

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES  
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum  
Internationales Büro



(43) Internationales Veröffentlichungsdatum  
17. März 2005 (17.03.2005)

PCT

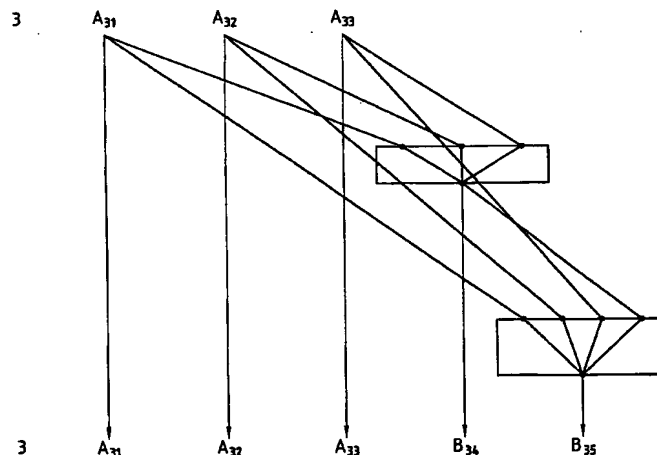
(10) Internationale Veröffentlichungsnummer  
**WO 2005/024717 A1**

- (51) Internationale Patentklassifikation<sup>7</sup>: **G06N 3/04**, 3/08, G06F 17/60
- (21) Internationales Aktenzeichen: PCT/CH2003/000612
- (22) Internationales Anmeldedatum:  
10. September 2003 (10.09.2003)
- (25) Einreichungssprache: Deutsch
- (26) Veröffentlichungssprache: Deutsch
- (71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von US): **SWISS REINSURANCE COMPANY [CH/CH]**; Mythenquai 60, CH-8002 Zürich (CH).
- (72) Erfinder; und
- (75) Erfinder/Anmelder (nur für US): **CUYPERS, Frank** [BE/CH]; Grossmannstrasse 30, CH-8049 Zürich (CH).
- (74) Anwalt: **BOVARD AG**; Optingenstrasse 16, CH-3000 Bern 25 (CH).
- (81) Bestimmungsstaaten (national): AE, AG, AL, AM, AT (Gebrauchsmuster), AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ (Gebrauchsmuster), DE (Gebrauchsmuster), DK (Gebrauchsmuster), DM, DZ, EC, EE (Gebrauchsmuster), EE, EG, ES, FI (Gebrauchsmuster), FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK (Gebrauchsmuster), SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Bestimmungsstaaten (regional): ARIPO Patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Fortsetzung auf der nächsten Seite]

(54) Title: SYSTEM AND METHOD FOR THE AUTOMATED ESTABLISHMENT OF EXPERIENCE RATINGS AND/OR RISK RESERVES

(54) Bezeichnung: SYSTEM UND VERFAHREN ZUR AUTOMATISIERTEN ERFAHRUNGSTARIFIERUNG UND/ODER SCHADENSRESERVIERUNG



(57) Abstract: The invention relates to a method for the automated establishment of experience ratings and/or risk reserves of events, whereby a certain event  $P_{i,f}$  of a starting year  $i$  includes development values  $P_{i,k}$  covering the development year  $k$ . For  $i, k$  it holds that  $i=1, \dots, K$  and  $k=1, \dots, K$ , with  $K$  being the last known development year and the first starting year  $i=1$  comprising all development values  $P_{i,k}$  in a defined manner. In order to determine the development values  $P_{i,k-(i-j)+1,f}$ , iterative  $(i-1)$  neuronal networks  $N_{i,j}$  are generated for every starting year  $i$ , with  $j=1, \dots, (i-1)$  being the number of iterations for a certain starting year  $i$  and the neuronal network  $N_{i,j+1}$  depending recursively on the neuronal network  $N_{i,j}$ . The inventive system and method is especially suitable for establishing experience ratings for insurance contracts and/or excess of loss reinsurance contracts.

[Fortsetzung auf der nächsten Seite]

WO 2005/024717 A1



**Erklärung gemäß Regel 4.17:**

— *Erfindererklärung (Regel 4.17 Ziffer iv) nur für US*

**Veröffentlicht:**

— *mit internationalem Recherchenbericht*

*Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.*

**(57) Zusammenfassung:** System sowie Verfahren zur automatisierten Erfahrungstarifizierung und/oder Schadensreservierung von Ereignissen, wobei ein bestimmtes Ereignis  $P_{i,f}$  eines Anfangsjahres  $i$  Entwicklungswerte  $P_{i,k}$  mit Entwicklungsjahr  $k$  umfasst. Für  $i, k$  gilt  $i=1, \dots, K$  und  $k=1, \dots, K$ , wobei  $K$  das letzt bekannte Entwicklungsjahr ist und wobei das erste Anfangsjahr  $i=1$  alle Entwicklungswerte  $P_{1,k}$  vorgegeben umfasst. Zur Bestimmung der Entwicklungswerte  $P_{i,K-(i-j)+1,f}$  werden für jedes Anfangsjahr  $i$  iterativ  $(i-1)$  neuronale Netzwerke  $N_{i,j}$  erzeugt, wobei die Anzahl Iterationen für ein bestimmtes Anfangsjahr  $i$  sind und wobei das neuronale Netzwerk  $N_{i,j+1}$  rekursiv vom neuronalen Netzwerk  $N_{i,j}$  abhängt. Insbesondere eignet sich das System und Verfahren zur Erfahrungstarifizierung von Versicherungsverträgen und/oder Schadenexzendenten-Rückversicherungsverträgen.

## **System und Verfahren zur automatisierten Erfahrungstarifizierung und/oder Schadensreservierung**

Die Erfindung betrifft ein System und ein Verfahren zur automatisierten Erfahrungstarifizierung und/oder Schadensreservierung, wobei  
5 ein bestimmtes Ereignis  $P_{if}$  eines Anfangszeitintervalles  $i$  mit  $f=1, \dots, F_i$  für eine Folge von Entwicklungsintervallen  $k=1, \dots, K$  Entwicklungswerte  $P_{ikf}$  umfasst. Für die Ereignisse  $P_{if}$  des ersten Anfangszeitintervall  $i=1$  sind alle Entwicklungswerte  $P_{1kf}$   $f=1, \dots, F_1$  bekannt. Die Erfindung betrifft insbesondere ein Computerprogrammprodukt zur Durchführung dieses Verfahrens.

10 Erfahrungstarifizierung bezieht sich im Stand der Technik auf Werteentwicklungen von Parametern von Ereignissen, die sich in einem bestimmten Jahr, dem Anfallsjahr oder Anfangsjahr, zum ersten Mal ereignen, und deren Folgen sich über mehrere Jahre fortpflanzen, den sog. Entwicklungsjahren. Allgemeiner ausgedrückt ereignen sich die Ereignisse zu  
15 einem bestimmten Zeitpunkt und entwickeln sich in vorgegebenen Zeitintervallen. Die Ereigniswerte des gleichen Ereignisses zeigen dabei über die verschiedenen Entwicklungsjahre oder Entwicklungszeitintervalle eine abhängige zeitrückbezogene Entwicklung. Die Erfahrungstarifizierung der Werte findet durch Extrapolation bzw. den Vergleich mit der Werteentwicklung  
20 bekannter ähnlicher, zeitlich zurückliegender Ereignisse statt.

Ein typisches Beispiel im Stand der Technik ist die mehrjährige Erfahrungstarifizierung anhand von Schadenfällen, z.B. des Zahlungsstandes  $Z$  oder des Reservestandes  $R$  eines Schadensfalles bei Versicherungsgesellschaften oder Rückversicherern. Bei der  
25 Erfahrungstarifizierung von Schadenfällen kennt ein Versicherungsunternehmen die Entwicklung jedes einzelnen Schadensfalls vom Zeitpunkt der Schadensmeldung bis zum aktuellen Stand oder bis zur Regulierung. Bei der Erfahrungstarifizierung liegt die Begründung der klassischen Credibility-Formel durch ein stochastisches Modell etwa 30 Jahre zurück; seitdem sind zahlreiche  
30 Varianten des Modells entwickelt worden, so dass man heute von einer eigentlichen Credibility-Theorie sprechen kann. Das Hauptproblem bei der Anwendung von Credibility-Formeln bilden die unbekannten Parameter, die

durch die Struktur des Bestandes bestimmt sind. Als Alternative zu bekannten Schätzverfahren bietet sich im Stand der Technik z.B. auch ein spieltheoretischer Ansatz an: Der Aktuar oder Versicherungsmathematiker kennt Schranken für die Parameter und bestimmt die optimale Prämie für den ungünstigsten Fall. Die Credibility-Theorie umfasst auch eine Reihe von Modellen zur Reservierung für Spätschäden. Dabei gibt es eine Vielfalt von Reservierungsverfahren, die, anders als die Credibility-Formel, nicht von unbekannten Parametern abhängen. Auch hier umfasst der Stand der Technik Verfahren durch stochastische Modelle, die die Erzeugung der Daten beschreiben. Eine Reihe von Ergebnissen liegt vor allem für das Chain-Ladder-Verfahren als eines der bekanntesten Verfahren zum Berechnen von ausstehenden Zahlungsforderungen bzw. zur Extrapolation der Schadensfälle vor. Die Stärken des Chain-Ladder-Verfahrens liegen einerseits in seiner Einfachheit andererseits darin, dass das Verfahren annähernd verteilungsfrei ist, d.h., das Verfahren basiert auf beinahe keinen Annahmen. Verteilungsfreie oder nicht parametrische Verfahren sind besonders für Fälle geeignet, in welchen der Anwender nur ungenügend oder gar keine Angaben über die zu erwartende Verteilung (z.B. Gaussverteilung etc.) der zu entwickelnden Parameter machen kann.

Das Chain-Ladder-Verfahren bedeutet, dass von einem Ereignis oder Schaden  $P_{if}$  mit  $f=1, 2, \dots, F_i$  aus Anfalljahr  $i=1, \dots, I$  Werte  $P_{ikf}$  bekannt sind, wobei  $P_{ikf}$  z.B. den Zahlungsstand oder den Reservestand zum Ende jedes Abwicklungsjahres  $k=1, \dots, K$  sein kann. Ein Ereignis  $P_{if}$  besteht in diesem Fall also in einer Folge von Punkten

$$P_{if} = (P_{i1f}, P_{i2f}, \dots, P_{iKf})$$

von der die ersten  $K+1-i$  Punkte bekannt sind und die noch unbekannten Punkte  $(P_{i,K+2-i,f}, \dots, P_{i,K,f})$ , prognostiziert werden sollen. Die Werte der Ereignisse  $P_{if}$  bilden ein sog. Schadensdreieck oder allgemeiner ein Ereigniswertedreieck

$$\begin{pmatrix}
 P_{11f=1..F_1} & P_{12f=1..F_1} & P_{13f=1..F_1} & P_{14f=1..F_1} & P_{15f=1..F_1} \\
 P_{21f=1..F_2} & P_{22f=1..F_2} & P_{23f=1..F_2} & P_{24f=1..F_2} & \\
 P_{31f=1..F_3} & P_{32f=1..F_3} & P_{33f=1..F_3} & & \\
 P_{41f=1..F_4} & P_{42f=1..F_4} & & & \\
 P_{51f=1..F_5} & & & & 
 \end{pmatrix}$$

- Die Zeilen und Spalten werden gebildet durch die Schadensanfalljahre und die Abwicklungsjahre. Allgemein ausgedrückt stehen in den Zeilen z.B. die Anfangsjahre und in den Spalten die Entwicklungsjahre der untersuchten Ereignisse, wobei die Darstellung auch verschieden dazu sein kann. Das Chain-Ladder-Verfahren basiert nun auf den kumulierten Schadensdreiecken, deren Einträge  $C_{ij}$  z.B. entweder reine Schadenszahlungen oder Schadensaufwendungen (Schadenszahlungen plus Veränderung der Schadensreserven) sind. Für die kumulierten Matrixelemente  $C_{ij}$  gilt

$$C_{ij} = \sum_{f=1}^{F_i} P_{if}$$

damit folgt

$$\begin{pmatrix}
 \sum_{f=1}^{F_1} P_{11f} & \sum_{f=1}^{F_1} P_{12f} & \sum_{f=1}^{F_1} P_{13f} & \sum_{f=1}^{F_1} P_{14f} & \sum_{f=1}^{F_1} P_{15f} \\
 \sum_{f=1}^{F_2} P_{21f} & \sum_{f=1}^{F_2} P_{22f} & \sum_{f=1}^{F_2} P_{23f} & \sum_{f=1}^{F_2} P_{24f} & \\
 \sum_{f=1}^{F_3} P_{31f} & \sum_{f=1}^{F_3} P_{32f} & \sum_{f=1}^{F_3} P_{33f} & & \\
 \sum_{f=1}^{F_4} P_{41f} & \sum_{f=1}^{F_4} P_{42f} & & & \\
 \sum_{f=1}^{F_5} P_{51f} & & & & 
 \end{pmatrix}$$

- Aus den mit dem Chain-Ladder-Verfahren interpolierten kumulierten Werten kann auch wieder auf das Einzelereignis geschlossen werden, indem eine Bestimmte Verteilung, z.B. typischerweise eine Pareto-Verteilung, der Werte angenommen wird. Die Pareto-Verteilung ist insbesondere für Versicherungsarten wie z.B. Versicherungen von Grossschäden oder Rückversicherer etc. geeignet. Die Pareto-Verteilung hat folgende Form

$$\Theta(x) = 1 - \left(\frac{x}{T}\right)^\alpha$$

wobei T ein Schwellwert und  $\alpha$  der Fitparameter ist. Die Einfachheit des Chain-Ladder-Verfahrens liegt insbesondere darin, dass es zur Anwendung nicht mehr als obiges (über den Entwicklungswerten der einzelnen Ereignisse kumuliertes) Schadensdreieck und z.B. keine Informationen über Meldedaten, Reservierungsabwicklungen oder Annahmen über mögliche Schadenshöhenverteilungen etc. braucht. Die Nachteile des Chain-Ladder-Verfahrens sind im Stand der Technik hinlänglich bekannt (siehe z.B. Thomas Mack, Measuring the Variability of Chain Ladder Reserve Estimates, submitted CAS Prize Paper Competition 1993, Greg Taylor, Chain Ladder Bias, Centre for Actuarial Studies, University of Melbourne, Australia, March 2001, pp 3). Um einen guten Schätzwert zu erhalten ist eine ausreichende Datenhistorie notwendig. Insbesondere bewährt sich das Chain Ladder Verfahren bei Sparten, wie z.B. Kfz-Haftpflichtversicherungen, bei welchen die Unterschiede in den Schadensjahren zu grossen Teilen auf Unterschiede in den Schadenfrequenzen zurückzuführen sind, da die Schätzer des Chain-Ladder-Verfahrens den Maximum-Likelihood Schätzern eines Modells mittels modifizierter Poissonverteilung entsprechen. Vorsicht ist daher z.B. bei Jahren geboten, in denen Änderungen an der Schadenshöhenverteilung (z.B. eine Erhöhung der Höchsthaftungssumme oder Änderungen im Selbstbehalt) vorgenommen werden, da diese Änderungen zu Strukturbrüchen im Chain-Ladder-Verfahren führen können. Bei Sparten mit extrem langer Abwicklungsdauer - wie z.B. bei der allgemeinen Haftpflichtversicherung - führt in vielen Fällen die Anwendung des Chain-Ladder-Verfahrens ebenfalls zu brauchbaren Ergebnissen, obwohl Informationen, wie z.B. eine zuverlässige Schätzung der Endschadensquote, aufgrund der langen Abwicklungsdauer selten verfügbar sind. Der Hauptnachteil des Chain-Ladder-Verfahrens liegt jedoch darin, dass das Chain-Ladder-Verfahren auf dem kumulierten Schadensdreieck basiert, d.h. durch die Kumulierung der Ereigniswerte der Ereignisse mit gleichem Anfangsjahr geht wesentliche Information über die einzelnen Schäden bzw. Ereignisse verloren, die später nicht mehr zurückerhalten werden kann.

Im Stand der Technik ist ein Verfahren von T. Mack bekannt (Thomas Mack, Schriftreihe Angewandte Versicherungsmathematik, Heft 28, p 310ff, Verlag Versicherungswirtschaft E.V., Karlsruhe 1997), bei welchem die Werte fortgepflanzt werden können, d.h. die Werte im Schadensdreieck  
 5 extrapoliert werden können, ohne dass die Informationen der Einzelereignisse verloren geht. Mit dem Mack'schen Verfahren kann also unter Ausnutzung der vollständigen Zahlenbasis für jeden Schaden eine individuelle IBNER-Reserve berechnet werden (IBNER: Incurred But Not Enough Reported). Unter IBNER-Forderungen versteht man Zahlungsforderungen, welche entweder über den  
 10 prognostizierten Werten liegen oder noch ausstehend sind. Die IBNER-Reserve ist insbesondere zur Erfahrungstarifizierung von Schadenexzedenten-Rückversicherungsverträgen nützlich, wo der Rückversicherer in der Regel zumindest für die relevanten Grossschäden die erforderlichen Einzelschadensdaten erhält. Beim Rückversicherer beschreibt die zeitliche  
 15 Entwicklung eines Bestandes von Risiken durch einen Risiko-Prozess, in dem Schadenzahlen und Schadenhöhen modelliert werden, wodurch in der Schadenexzedenten-Rückversicherung beim Übergang vom Erst- zum Rückversicherer das Phänomen der zufälligen Verdünnung des Risiko-Prozesses entsteht; andererseits werden durch Rückversicherung Bestände  
 20 von mehreren Erstversicherern zusammengeführt und damit Risiko-Prozesse überlagert. Die Auswirkungen von Verdünnung und Überlagerung wurden bisher vor allem für Poisson'sche Risiko-Prozesse untersucht. Für Versicherungen/Rückversicherungen bedeutet die Erfahrungstarifizierung mittels des Mack'schen Verfahrens, dass von jedem Schaden  $P_{if}$  mit  $f=1, 2, \dots, F_i$  aus  
 25 Anfalljahr oder Anfangsjahr  $i=1, \dots, I$  der Zahlungsstand  $Z_{ikf}$  und der Reservestand  $R_{ikf}$  zum Ende jedes Abwicklungsjahres oder Entwicklungsjahres  $k=1, \dots, K$  bis zum aktuellen Stand  $(Z_{i,K+1-i,f}, R_{i,K+1-i,f})$  bekannt ist. Ein Schaden  $P_{if}$  besteht in diesem Fall also in einer Folge von Punkten

$$P_{if} = (Z_{i1f}, R_{i1f}), (Z_{i2f}, R_{i2f}), \dots, (Z_{iKf}, R_{iKf})$$

30 in der Zahlungs-Reserve-Ebene, von der die ersten  $K+1-i$  Punkte bekannt sind und die noch unbekannten Punkte  $(Z_{i,K+2-i,f}, R_{i,K+2-i,f}), \dots, (Z_{i,K,f}, R_{i,K,f})$ , prognostiziert werden sollen. Insbesondere interessiert natürlich der Endstand  $(Z_{i,K,f}, R_{i,K,f})$ , wobei im Idealfall  $R_{i,K,f} = 0$  ist, d.h. der Schaden als

vollständig reguliert angesehen wird; ob dies erreichbar ist, hängt von der Länge K des betrachteten Entwicklungszeitraums ab. Ein Schadensstand  $(Z_{i,K+1-i,f}, R_{i,K+1-i,f})$  wird im Stand der Technik, wie z.B. im Mack'schen Verfahren, so fortgesetzt, wie dies bei ähnlichen Schäden aus früheren Anfalljahren der Fall war. In den herkömmlichen Verfahren muss also zum einen festgelegt werden, wann zwei Schäden "ähnlich" sind und zum anderen, was es heisst einen Schaden "fortzusetzen". Ausserdem muss neben der sich so ergebenden IBNER-Reserve in einem zweiten Schritt festgelegt werden, wie die echten Spätschäden zu berechnen sind, von denen zum aktuellen Zeitpunkt noch nichts bekannt ist.

Zur Qualifizierung der Ähnlichkeit wird im Stand der Technik z.B. der Euklidische Abstand

$$d((Z, R), (\tilde{Z}, \tilde{R})) = \sqrt{(Z - \tilde{Z})^2 + (R - \tilde{R})^2}$$

in der Zahlungs-Reserve-Ebene verwendet. Aber auch mit dem Euklidischen Abstand gibt es viele Möglichkeiten, um zu einem gegebenen Schaden  $(P_{i,1,f}, P_{i,2,f}, \dots, P_{i,K+1-i,f})$  den nächst ähnlichsten Schaden eines früheren Anfalljahrs zu finden, d.h. den Schaden  $(\tilde{P}_1, \dots, \tilde{P}_k)$  mit  $k > K+1-i$ , für den entweder

$$\sum_{j=1}^{K+1-i} d(P_{ijf}, \tilde{P}_j) \quad (\text{Summe aller bisherigen Abstände})$$

20

oder

$$\sum_{j=1}^{K+1-i} j \cdot d(P_{ijf}, \tilde{P}_j) \quad (\text{gewichtete Summe aller Abstände})$$

oder

$$\max_{1 \leq j \leq K+1-i} d(P_{ijf}, \tilde{P}_j) \quad (\text{maximaler Abstand})$$

oder



$$d(P_{i,K+1-i,f}, \tilde{P}_{K+1-j}) \quad (\text{aktueller Abstand})$$

minimal ist.

Beim Beispiel des Mack'schen Verfahrens wird im Normalfall der aktuelle Abstand benutzt. Dies bedeutet, dass zu einem Schaden  $(P_1, \dots, P_k)$ , dessen Abwicklung bis zum k-ten Entwicklungsjahr bekannt ist, von allen anderen Schäden  $(\tilde{P}_1, \dots, \tilde{P}_j)$ , deren Entwicklung mindestens bis Entwicklungsjahr  $j \geq k + 1$  bekannt ist, derjenige als der Ähnlichste betrachtet wird, für den der aktuelle Abstand  $d(P_k, \tilde{P}_k)$  am kleinsten ist.

Der Schaden  $(P_1, \dots, P_k)$  wird nun so fortgesetzt, wie dies bei seinem abstandsnächsten „Vorbild“  $(\tilde{P}_1, \dots, \tilde{P}_k, \tilde{P}_{k+1}, \dots, \tilde{P}_j)$  der Fall ist. Hierfür gibt es die Möglichkeit, um ein einziges Abwicklungsjahr (d.h. bis  $P_{k+1}$ ) oder um mehrere Entwicklungsjahre zugleich (z.B. bis  $P_j$ ) fortzusetzen. Bei Verfahren, wie z.B. dem Mack'schen-Verfahren wird typischerweise zunächst nur um ein Abwicklungsjahr fortgesetzt, um dann wieder einen neuen ähnlichsten Schaden zu suchen, womit der gerade fortgesetzte Schaden um ein weiteres Entwicklungsjahr fortgesetzt wird. Der nächste gefundene Schaden kann natürlich auch wieder derselbe sein. Zur Fortsetzung der Schadensfälle gibt es zwei Möglichkeiten. Die additive Fortsetzung von  $P_k = (Z_k, R_k)$

$$\hat{P}_{k+1} = (\hat{Z}_{k+1}, \hat{R}_{k+1}) = (Z_k + \tilde{Z}_{k+1} - \tilde{Z}_k, R_k + \tilde{R}_{k+1} - \tilde{R}_k),$$

und die multiplikative Fortsetzung von  $P_k = (Z_k, R_k)$

$$\hat{P}_{k+1} = (\hat{Z}_{k+1}, \hat{R}_{k+1}) = (Z_k \cdot \frac{\tilde{Z}_{k+1}}{\tilde{Z}_k}, R_k \cdot \frac{\tilde{R}_{k+1}}{\tilde{R}_k}).$$

Es ist einfach zu sehen, dass einer der Nachteile des Standes der Technik, insbesondere des Mack'schen Verfahrens, u.a. in der Art der Fortsetzung der Schadensfälle liegt. Die multiplikative Fortsetzung ist nur für so genannte offene Schadensstände, d.h.  $Z_k > 0$ ,  $R_k > 0$  sinnvoll. Bei mutmasslichen Schadensständen  $P_k = (0, R_k)$ ,  $R_k > 0$ , muss die multiplikative Fortsetzung diversifiziert werden, da andernfalls keine Fortsetzung erfolgt. Ausserdem, falls  $\tilde{Z}_k = 0$  oder  $\tilde{R}_k = 0$ , erfolgt eine Division durch 0. Ähnlich,

falls  $\tilde{Z}_k$  oder  $\tilde{R}_k$  klein ist, kann das multiplikative Verfahren leicht zu unrealistisch hohen Fortsetzungen führen. Dies lässt eine konsistente Behandlung der Fälle nicht zu. D.h. die Reserve  $R_k$  kann in diesem Fall nicht einfach fortgesetzt werden. Ebenso kann ein ausregulierter Schadensstand  $P_k = (Z_k, 0)$ ,  $Z_k > 0$  ebenfalls nicht weiterentwickelt werden. Eine Möglichkeit ist, ihn einfach unverändert zu lassen. Ein Wiederaufleben eines Schadens wird damit jedoch verhindert. Allenfalls könnte man ihn anhand des nächstliegenden ausregulierten Vorbildes fortsetzen, was ebenfalls eine konsistente Behandlung der Fälle nicht zulässt. Auch bei der additiven Fortsetzung sollten sinnvollerweise mutmassliche Schadensstände nur anhand eines ebenfalls mutmasslichen Vorbildes fortgesetzt werden, um den Euklidischen Abstand zu minimieren und um eine entsprechende Qualifizierung der Ähnlichkeit zu garantieren. Ein analoger Nachteil entsteht bei ausregulierten Schadensständen, falls ein Wiederaufleben zugelassen werden soll und negative Reserven vermieden werden sollen. Ganz allgemein kann das additive Verfahren leicht zu negativen Zahlungen und/oder Reserven führen. Zusätzlich kann im Stand der Technik ein Schaden  $P_k$  nicht fortgesetzt werden, wenn kein entsprechendes Vorbild existiert, ohne dass nicht weitere Annahmen in das Verfahren gesteckt werden. Als Beispiel dazu ist ein offener Schaden  $P_k$ , wenn es im gleichen Abwicklungsjahr  $k$  keinen Schaden aus früheren Anfallsjahren gibt, bei dem  $\tilde{P}_k$  ebenfalls offen ist. Ein Ausweg aus dem Dilemma kann dadurch gefunden werden, dass für diesen Fall  $P_k$  unverändert gelassen wird, d.h.  $\hat{P}_{k+1} = P_k$ , was natürlich keiner echten Fortsetzung entspricht.

Insgesamt wird somit im Stand der Technik jeder aktuelle Schadensstand  $P_{i,K+1-i,f} = (Z_{i,K+1-i,f}, R_{i,K+1-i,f})$  schrittweise bis zum Entwicklungs- bzw. Abwicklungsende nach  $K$ -Entwicklungsjahren entweder additiv oder multiplikativ weiterentwickelt. Dabei wird in jedem Schritt der jeweils gemäss Euklidischem Abstand nächstliegende Vorbildschadensstand vom gleichen Schadensstandstyp (mutmasslich, offen oder ausreguliert) festgestellt und der fortzusetzende Schadensstand entsprechend der Weiterentwicklung des Vorbildschadens entweder additiv oder multiplikativ fortgesetzt. Für das Mack'sche Verfahren ist es ebenfalls sinnvoll, als Vorbild stets nur tatsächlich beobachtete Schadensentwicklungen  $\tilde{P}_k \rightarrow \tilde{P}_{k+1}$  in Betracht zu ziehen und keine extrapolierte, d.h. entwickelten Schadenentwicklungen, da andernfalls eine

Korrelation und/oder ein entsprechender Bias der Ereignisse nicht zu vermeiden ist. Umgekehrt erhält man damit jedoch den Nachteil, dass bereits bekannte Information von Ereignissen verloren geht.

Von der Konstruktion der Verfahren des Standes der Technik  
5 leuchtet es unmittelbar ein, dass die Verfahren auch getrennt, einerseits auf das Dreieck der Zahlungen, andererseits auf das Dreieck der Reserven, angewandt werden können. Natürlich könnten bei der beschriebenen Vorgehensweise auch andere Möglichkeiten zugelassen werden, um den jeweils nächstliegenden Schadensstand als Vorbild zu finden. Dies würde sich  
10 jedoch insbesondere auf die Verteilungsfreiheit des Verfahrens auswirken. Damit lässt sich sagen, dass sich beim Standes der Technik die oben genannten systematischen Probleme auch durch entsprechende Modifikationen nicht beheben lassen oder allenfalls nur dadurch, dass weitere Modellannahmen in das Verfahren gesteckt werden. Gerade aber bei  
15 komplexen dynamisch nichtlinearen Prozessen, wie z.B. die Entwicklung von Schadensfällen, ist dies in den aller meisten Fällen nicht wünschenswert. Selbst wenn man die genannten Nachteile ausser Acht lässt, muss bei dem herkömmlichen Verfahren nach T. Mack immer noch festgelegt werden, wann zwei Schäden ähnlich sind und was es heisst, einen Schaden fortzusetzen,  
20 wodurch also minimale Grundannahmen bzw. Modelannahmen getroffen werden müssen. Im Stand der Technik ist jedoch nicht nur die Wahl der euklidischen Metrik arbiträr, sondern auch die Wahl zwischen dem erwähnten multiplikativen und additiven Verfahren. Weiter wird im Stand der Technik die Fehlerschätzung nicht näher definiert. Zwar ist es vorstellbar, einen Fehler z.B.  
25 basierend auf der inversen Distanz zu definieren. Dies wird aber im Stand der Technik nicht offenbart. Ein wichtiger Nachteil des Standes der Technik ist jedoch auch, dass jedes Ereignis mit allen vorherigen verglichen werden muss, um fortgesetzt werden zu können. Der Aufwand steigt linear mit der Anzahl Jahre und linear mit der Anzahl Schäden im Portefeuille. Wenn Portefeuilles  
30 aggregiert, steigt der Rechenaufwand und der Speicherbedarf entsprechend.

Neuronale Netze sind grundsätzlich im Stand der Technik bekannt und werden z.B. zum Lösen von Optimierungsaufgaben, Bildererkennung (Patternrecognition), in der künstlichen Intelligenz etc. eingesetzt. Entsprechend

biologischer Nervenetze besteht ein neuronales Netzwerk aus einer Vielzahl von Netzknoten, sog. Neuronen, die über gewichtete Verbindungen (Synapsen) miteinander verbunden sind. Die Neuronen sind in Netzsichten (Layers) organisiert und zusammengeschaltet. Die einzelnen Neuronen werden in

5 Abhängigkeit ihrer Eingangssignale aktiviert und erzeugen ein entsprechendes Ausgangssignal. Die Aktivierung eines Neurons erfolgt über einen individuellen Gewichtungsfaktor durch die Summation über die Eingangssignale. Derartige neuronale Netze sind lernfähig, indem die Gewichtungsfaktoren in Abhängigkeit von vorgegebenen beispielhaften Eingangs- und Ausgangswerten systematisch

10 so lange verändert werden, bis das neuronale Netz in einem definierten vorhersagbaren Fehlerbereich ein gewünschtes Verhalten zeigt, wie z. B. die Vorhersage von Ausgabewerten für zukünftige Eingangswerte. Damit weisen neuronale Netze adaptive Fähigkeiten zur Erlernung und Speicherung von Wissen und assoziative Fähigkeiten zum Vergleich von neuen Informationen

15 mit gespeichertem Wissen auf. Die Neuronen (Netzknoten) können einen Ruhezustand oder einen Erregungszustand einnehmen. Jedes Neuron hat mehrere Eingänge und genau einen Ausgang, der mit den Eingängen anderer Neuronen der nachfolgenden Netzsicht verbunden ist oder im Falle eines Ausgangsknotens einen entsprechenden Ausgangswert repräsentiert. Ein

20 Neuron geht in den Erregungszustand über, wenn eine genügende Anzahl der Eingänge des Neurons über einem bestimmten Schwellenwert des Neurons erregt sind, d.h. falls die Summation über den Eingängen einen bestimmten Schwellenwert erreicht. In den Gewichten der Eingänge eines Neurons und in dem Schwellenwert des Neurons ist das Wissen durch Adaption abgespeichert.

25 Mittels Lernvorgang werden die Gewichte eines neuronalen Netzes trainiert (siehe z.B. G. Cybenko, "Approximation by Superpositions of a sigmoidal function", Math. Control, Sig. Syst., 2, 1989, pp 303-314; M.T. Hagan, M.B. Menhaj, "Training Feedforward Networks with the Marquardt Algorithm", IEEE Transactions on Neural Networks, Vol. 5, Nr. 6, pp 989-993, November 1994;

30 K. Hornik, M. Stinchcombe, H. White, "Multilayer Feedforward Networks are universal Approximators", Neural Networks, 2, 1989, pp 359-366 etc.).

Es ist eine Aufgabe dieser Erfindung, ein neues System und Verfahren zur automatisierten Erfahrungstarifizierung von Ereignissen und/oder Schadensreservierung vorzuschlagen, das die oben genannten Nachteile des

Standes der Technik nicht aufweist. Insbesondere soll ein automatisiertes, einfaches und rationelles Verfahren vorgeschlagen werden, um einen vorgegebenen Schaden mit einem individuellen Zuwachs bzw. Faktor weiterzuentwickeln, so dass nachträglich die gesamte Information über die  
5 Entwicklung eines einzelnen Schadens zur Verfügung steht. Bei dem Verfahren sollen so wenig wie möglich Annahmen über die Verteilung von vornherein getroffen werden und gleichzeitig soll die maximal mögliche Information der vorgegebenen Fälle ausgenutzt werden.

Gemäss der vorliegenden Erfindung wird dieses Ziel insbesondere  
10 durch die Elemente der unabhängigen Ansprüche erreicht. Weitere vorteilhafte Ausführungsformen gehen ausserdem aus den abhängigen Ansprüchen und der Beschreibung hervor.

Insbesondere werden diese Ziele durch die Erfindung dadurch erreicht, dass einem bestimmten Ereignis  $P_{i,f}$  eines Anfangszeitintervalles  $i$   
15 Entwicklungswerte  $P_{i,k,f}$  mit Entwicklungsintervallen  $k=1,...,K$  zugeordnet werden, bei welchem  $K$  das letzte bekannte Entwicklungsintervall ist mit  $i=1,...,K$  und für die Ereignisse  $P_{i,f}$  alle Entwicklungswerte  $P_{i,k,f}$  bekannt sind, wobei zur Bestimmung der Entwicklungswerte  $P_{i,k+2,f}, \dots, P_{i,K,f}$  mindestens ein neuronales Netzwerk verwendet wird. Bei bestimmten Ereignissen kann z.B. das  
20 Anfangszeitintervall einem Anfangsjahr zugeordnet werden und die Entwicklungsintervalle Entwicklungsjahren zugeordnet werden. Die Entwicklungswerte  $P_{i,k,f}$  der verschiedenen Ereignisse  $P_{i,f}$  können entsprechend ihrem Anfangszeitintervall mittels mindestens einem Normierungsfaktor normiert werden. Die Normierung der Entwicklungswerte  $P_{i,k,f}$  hat u.a. den  
25 Vorteil, dass die Entwicklungswerte zu unterschiedlichen Zeitpunkten vergleichbar sind. Diese Ausführungsvariante hat weiter u.a. den Vorteil, dass zur automatisierten Erfahrungstarifizierung keine Modelannahmen z.B. über Werteverteilungen, Systemdynamiken etc. vorausgesetzt werden müssen. Insbesondere ist die Erfahrungstarifizierung frei von  
30 Proximationsvoraussetzungen wie z.B. das Euklidische Mass etc. Dies ist beim Stand der Technik so nicht möglich. Zusätzlich wird die gesamte Information des Datensamples verwendet, ohne dass die Datensätze kumuliert werden. Die komplette Information über die einzelnen Ereignisse bleibt in jeder Stufe

erhalten und kann am Ende wieder abgerufen werden. Die Normierung hat den Vorteil, dass Datensätze unterschiedlicher Anfangszeitintervalle vergleichbare Größenordnungen erhalten und so besser verglichen werden können.

In einer Ausführungsvariante werden zur Bestimmung der

5 Entwicklungswerte  $P_{i,K-(i-j)+1,f}$  iterativ  $(i-1)$  neuronale Netzwerke  $N_{i,j}$  für jedes Anfangszeitintervall und/oder Anfangsjahr  $i$  erzeugt mit  $j=1, \dots, (i-1)$ , wobei das neuronale Netzwerk  $N_{i,j+1}$  rekursiv vom neuronalen Netzwerk  $N_{i,j}$  abhängt. Zum Gewichten eines bestimmten neuronalen Netzwerkes  $N_{i,j}$  können z.B. die Entwicklungswerte  $P_{p,q,f}$  mit  $p=1, \dots, (i-1)$  und  $q=1, \dots, K-(i-j)$  verwendet werden.

10 Diese Ausführungsvariante hat u.a. den Vorteil, dass wie in der vorhergehenden Ausführungsvariante die gesamte Information des Datensamples verwendet wird, ohne dass die Datensätze kumuliert werden. Die komplette Information über die einzelnen Ereignisse bleibt in jeder Stufe erhalten und kann am Ende wieder abgerufen werden. Mittels einer

15 Minimierung eines global eingeführten Fehlers können die Netzwerke zusätzlich optimiert werden.

In einer anderen Ausführungsvariante werden die neuronalen

Netzwerk  $N_{i,j}$  für gleiche Entwicklungsjahre und/oder Entwicklungsintervalle  $j$  identisch trainiert, wobei für ein Anfangszeitintervall und/oder Anfangsjahr  $i+1$

20 das neuronale Netzwerk  $N_{i+1,j=i}$  erzeugt wird und alle anderen neuronalen Netzwerke  $N_{i+1,j<i}$  von früheren Anfangszeitintervallen und/oder Anfangsjahren übernommen werden. Diese Ausführungsvariante hat u.a. den Vorteil, dass nur bekannte Daten zur Erfahrungstarifizierung verwendet werden und vom System bestimmte Daten nicht weiter verwendet werden, wodurch die Korrelation der

25 Fehler bzw. der Daten verhindert wird.

In einer wieder anderen Ausführungsvariante werden zur Bestimmung zusätzlich Ereignisse  $P_{i,f}$  mit Anfangszeitintervall  $i<1$  verwendet werden, wobei für die Ereignisse  $P_{i<1,f}$  alle Entwicklungswerte  $P_{i<1,k,f}$  bekannt sind. Diese Ausführungsvariante hat u.a. den Vorteil, dass durch die

30 zusätzlichen Datensätze die neuronalen Netzwerke besser optimiert werden können und ihr Fehler minimiert werden kann.

In einer weiteren Ausführungsvariante werden zur automatisierten Erfahrungstarifizierung und/oder Schadensreservierung einem bestimmten Ereignis  $P_{i,f}$  eines Anfangszeitintervalles  $i$  Entwicklungswerte  $P_{i,k,f}$  mit Entwicklungsintervallen  $k=1,\dots,K$  zugeordnet abgespeichert, bei welchem  $i =$

5  $1,\dots,K$  und  $K$  das letzte bekannte Entwicklungsintervall ist und bei welchem für das erste Anfangszeitintervall alle Entwicklungswerte  $P_{1,k,f}$  bekannt sind, wobei für jedes Anfangszeitintervall  $i=2,\dots,K$  mittels Iterationen  $j=1,\dots,(i-1)$  bei jeder Iteration  $j$  in einem ersten Schritt ein neuronales Netzwerk  $N_{i,j}$  mit einem Inputlayer mit  $K-(i-j)$  Inputsegmenten und einem Outputlayer erzeugt wird,

10 welche Inputsegmente mindestens ein Inputneuron umfassen und einem Entwicklungswert  $P_{i,k,f}$  zugeordnet werden, wobei in einem zweiten Schritt das neuronale Netzwerk  $N_{i,j}$  mit den verfügbaren Ereignissen  $P_{i,f}$  aller Anfangszeitintervalle  $m=1,\dots,(i-1)$  mittels der Entwicklungswerte  $P_{m,1..K-(i-j),f}$  als Input und  $P_{m,1..K-(i-j)+1,f}$  als Output gewichtet wird, und wobei in einem dritten

15 Schritt mittels des neuronalen Netzwerkes  $N_{i,j}$  die Outputwerte  $O_{i,f}$  für alle Ereignisse  $P_{i,f}$  des Anfangszeitintervalles  $i$  bestimmt werden, wobei der Outputwert  $O_{i,f}$  dem Entwicklungswert  $P_{i,K-(i-j)+1,f}$  des Ereignisses  $P_{i,f}$  zugeordnet wird und wobei das neuronale Netzwerk  $N_{i,j}$  rekursiv vom neuronalen Netzwerk  $N_{i,j+1}$  abhängt. Bei bestimmten Ereignissen kann z.B. das Anfangszeitintervall

20 einem Anfangsjahr zugeordnet werden und die Entwicklungsintervalle Entwicklungsjahren zugeordnet werden. Diese Ausführungsvariante hat u.a. die gleichen Vorteile wie die vorhergehenden Ausführungsvarianten.

In einer Ausführungsvariante umfasst ein System neuronale Netzwerke  $N_i$  mit jeweils einen Inputlayer mit mindestens einem Inputsegment

25 und einen Outputlayer, welcher Input- und Outputlayer eine Vielzahl von Neuronen umfasst, die gewichtet miteinander verbunden sind, wobei die neuronalen Netzwerke  $N_i$  mittels einer Recheneinheit software- und/oder hardwaremässig iterativ erzeugbar sind, wobei ein neuronales Netzwerk  $N_{i+1}$  rekursiv vom neuronalen Netzwerk  $N_i$  abhängt und jedes Netzwerk  $N_{i+1}$  jeweils

30 ein Inputsegment mehr als das Netzwerk  $N_i$  umfasst, wobei jedes neuronale Netzwerk  $N_i$  beginnend beim neuronalen Netzwerk  $N_1$  mittels eines Minimierungsmoduls durch Minimierung eines lokal propagierten Fehlers trainierbar ist, und wobei das rekursive System von neuronalen Netzwerken mittels einem Minimierungsmodul durch Minimierung eines global propagierten

Fehlers basierend auf den lokalen Fehlern der neuronalen Netzwerke  $N_i$  trainierbar ist. Diese Ausführungsvariante hat u.a. den Vorteil, dass die rekursiv erzeugten neuronalen Netzwerke mittels des globalen Fehlers zusätzlich optimiert werden können. U.a. ist es die Kombination der Rekursiven

- 5 Erzeugung der neuronalen Netzwerkstruktur mit einer doppelten Minimierung mittels lokal propagiertem Fehler und global propagierten Fehler, welches die Vorteile dieser Ausführungsvariante ergibt.

In einer anderen Ausführungsvariante ist der Outputlayer des neuronalen Netzwerkes  $N_i$  mit mindestens einem Inputsegment des Inputlayers  
10 des neuronalen Netzwerkes  $N_{i+1}$  zugeordnet verbunden. Diese Ausführungsvariante hat u.a. den Vorteil, dass das System von neuronalen Netzwerken wiederum als neuronales Netzwerk aufgefasst werden kann. Damit können Teilnetzwerke eines ganzen Netzwerkes lokal gewichtet werden und auch bei einem globalen Lernen durch das System mittels den entsprechenden  
15 Datensätzen in ihrem Verhalten kontrolliert und überwacht werden. Dies war bis anhin im Stand der Technik so nicht möglich.

An dieser Stelle soll festgehalten werden, dass sich die vorliegende Erfindung neben dem erfindungsgemässen Verfahren auch auf ein System zur Ausführung dieses Verfahrens bezieht. Ferner beschränkt es sich nicht auf das  
20 genannte System und Verfahren, sondern bezieht sich ebenso auf rekursiv geschachtelte Systeme von neuronalen Netzwerken und ein Computerprogrammprodukt zur Realisierung des erfindungsgemässen Verfahrens.

Nachfolgend werden Ausführungsvarianten der vorliegenden Erfindung  
25 anhand von Beispielen beschrieben. Die Beispiele der Ausführungen werden durch folgende beigelegte Figuren illustriert:

Figur 1 zeigt ein Blockdiagramm, welches schematisch die Trainings- bzw. Bestimmungsphase oder Präsentationsphase eines neuronalen Netzwerkes zur Bestimmung des Ereigniswertes  $P_{2,5,f}$  eines Ereignisses  $P_f$  in  
30 einer oberen 5x5 Matrix, d.h. bei  $K=5$ , wiedergibt. Die gestrichelte Linie T gibt



die Trainingsphase an und die ausgezogene Linie R die Bestimmungsphase nach dem Lernen.

Figur 2 zeigt ebenfalls ein Blockdiagramm, welches wie Fig. 1 schematisch die Trainings- bzw. Bestimmungsphase eines neuronalen Netzwerkes zur Bestimmung des Ereigniswertes  $P_{3,4,f}$  für das dritte Anfangsjahr wiedergibt.

Figur 3 zeigt ein Blockdiagramm, welches wie Fig. 1 schematisch die Trainings- bzw. Bestimmungsphase eines neuronalen Netzwerkes zur Bestimmung des Ereigniswertes  $P_{3,5,f}$  für das dritte Anfangsjahr wiedergibt.

Figur 4 zeigt ein Blockdiagramm, welches schematisch nur die Trainingsphase zur Bestimmung von  $P_{3,4,f}$  und  $P_{3,5,f}$  zeigt, wobei die berechneten Werte  $P_{3,4,f}$  zum Trainieren des Netzwerkes zur Bestimmung von  $P_{3,5,f}$  verwendet werden.

Figur 5 zeigt ein Blockdiagramm, welches schematisch die rekursive Erzeugung neuronaler Netzwerke zur Bestimmung der Werte in Zeile 3 einer 5x5 Matrix zeigt, wobei 2 Netzwerke erzeugt werden.

Figur 6 zeigt ein Blockdiagramm, welches schematisch die rekursive Erzeugung neuronaler Netzwerke zur Bestimmung der Werte in Zeile 5 einer 5x5 Matrix zeigt, wobei 4 Netzwerke erzeugt werden.

Figur 7 zeigt ein Blockdiagramm, welches schematisch ebenfalls ein erfindungsgemässes System zeigt, wobei die Trainingsbasis auf die bekannten Ereigniswerte  $A_{ij}$  eingeschränkt ist.

Figur 1 bis 7 illustrieren schematisch eine Architektur, die zur Realisierung der Erfindung verwendet werden kann. In diesem Ausführungsbeispiel umfasst zur automatisierten Erfahrungstarifizierung von Ereignissen und/oder Schadensreservierung ein bestimmtes Ereignis  $P_{i,f}$  eines Anfangsjahres  $i$  Entwicklungswerte  $P_{ikf}$ . Der Index  $f$  läuft über alle Ereignisse  $P_{i,f}$  für ein bestimmtes Anfangsjahr  $i$  mit  $f = 1, \dots, F_i$ . Der Entwicklungswert  $P_{ikf} = (Z_{ikf}$ ,

$R_{ikf}, \dots$ ) ist ein beliebiger Vektor und/oder n-Tupel von Entwicklungsparametern  $Z_{ikf}, R_{ikf}, \dots$ , welche für ein Ereignis entwickelt werden sollen. So kann z.B. bei Versicherungen für einen Schadensfall  $P_{ikf}$   $Z_{ikf}$  der Zahlungsstand sein,  $R_{ikf}$  der Reservestand, etc. Beliebige weitere relevante Parameter für ein Ereignis sind  
 5 vorstellbar, ohne dass dies den Schutzbereich der Erfindung berühren würde. Die Entwicklungsjahre  $k$  gehen aus von  $k=1, \dots, K$  und die Anfangsjahre  $i = 1, \dots, I$ .  $K$  ist das letzte bekannte Entwicklungsjahr. Für das erste Anfangsjahr  $i = 1$  sind alle Entwicklungswerte  $P_{1kf}$  vorgegeben. Wie bereits angegeben, sollen für dieses Beispiel die Anzahl Anfangsjahre  $I$  und die Anzahl Entwicklungsjahre  $K$   
 10 gleich sein, d.h.  $I = K$ . Es ist aber durchaus vorstellbar, dass  $I \neq K$ , ohne dass das Verfahren oder das System dadurch eingeschränkt würde.  $P_{ikf}$  ist also ein n-Tupel bestehend aus der Folge von Punkten und/oder Matrixelementen

$$(Z_{ikn}, R_{ikn}, \dots) \quad \text{mit } k = 1, 2, \dots, K$$

Mit  $I = K$  ergibt sich dadurch eine quadratische obere Dreiecksmatrix  
 15 und/oder Blockdreiecksmatrix für die bekannten Entwicklungswerte  $P_{ikf}$

$$\begin{pmatrix} P_{11f=1..F_1} & P_{12f=1..F_1} & P_{13f=1..F_1} & P_{14f=1..F_1} & P_{15f=1..F_1} \\ P_{21f=1..F_2} & P_{22f=1..F_2} & P_{23f=1..F_2} & P_{24f=1..F_2} & \\ P_{31f=1..F_3} & P_{32f=1..F_3} & P_{33f=1..F_3} & & \\ P_{41f=1..F_4} & P_{42f=1..F_4} & & & \\ P_{51f=1..F_5} & & & & \end{pmatrix}$$

wobei  $f=1, \dots, F_i$  wieder über alle Ereignisse für ein bestimmtes Anfangsjahr  $i$  geht. Damit sind die Zeilen der Matrix den Anfangsjahren und die Spalten der Matrix den Entwicklungsjahren zugeordnet. In dem  
 20 Ausführungsbeispiel soll  $P_{ikf}$  auf das Beispiel von Schadensfällen bei Versicherungen beschränkt werden, da insbesondere das Verfahren bzw. das System z.B. zur Erfahrungstarifizierung von Versicherungsverträgen und/oder Schadenexzendenten-Rückversicherungsverträgen sehr geeignet ist. Es muss betont werden, dass die Matrixelemente  $P_{ikf}$  selbst wieder Vektoren und/oder  
 25 Matrixen sein können, womit die obige Matrix zu einer entsprechenden Blockmatrix wird. Das erfindungsgemäße Verfahren und System eignet sich jedoch für Erfahrungstarifizierung bzw. zur Extrapolation von zeitverzögerten

nichtlinearen Prozessen ganz allgemein. Mit dem Gesagten ist  $P_{ikf}$  eine Folge von Punkten

$$(Z_{ikn}, R_{ikn}) \quad \text{mit } k = 1, 2, \dots, K$$

- in der Zahlungs-Reserve-Ebene, von der die ersten  $K+1-i$  Punkte  
 5 bekannt sind und die noch unbekannten Punkte  $(Z_{i,K+2-i,f}, R_{i,K+2-i,f}), \dots, (Z_{i,K,f}, R_{i,K,f})$ , prognostiziert werden sollen. Teilt man für diese Beispiel  $P_{ikf}$  nach Zahlungsebene und die Reserveebene auf, erhält man analog für die Zahlungsebene die Dreiecksmatrix

$$\begin{pmatrix} Z_{11f} & Z_{12f} & Z_{13f} & Z_{14f} & Z_{15f} \\ Z_{21f} & Z_{22f} & Z_{23f} & Z_{24f} & \\ Z_{31f} & Z_{32f} & Z_{33f} & & \\ Z_{41f} & Z_{42f} & & & \\ Z_{51f} & & & & \end{pmatrix}$$

10

und für die Reserveebene die Dreiecksmatrix

$$\begin{pmatrix} R_{11f} & R_{12f} & R_{13f} & R_{14f} & R_{15f} \\ R_{21f} & R_{22f} & R_{23f} & R_{24f} & \\ R_{31f} & R_{32f} & R_{33f} & & \\ R_{41f} & R_{42f} & & & \\ R_{51f} & & & & \end{pmatrix}$$

- Bei der Erfahrungstarifizierung von Schadensfällen ist somit die  
 15 Entwicklung jedes einzelnen Schadensfalls  $f_i$  vom Zeitpunkt der Schadensmeldung im Anfangsjahr  $i$  bis zum aktuellen Stand (aktuelles Entwicklungsjahr  $k$ ) oder bis zur Regulierung bekannt. Diese Information kann in einer Datenbank abgespeichert sein, welche Datenbank z.B. über ein Netzwerk mittels einer Recheneinheit abgerufen werden kann. Die Datenbank  
 20 kann jedoch auch direkt über einen internen Datenbus des erfindungsgemässen Systems zugreifbar sein oder anders ausgelesen werden.

- Um die Daten im Beispiel der Schadensfälle zu verwenden, werden die Dreiecksmatrizen in einem ersten Schritt normiert, d.h. die Schadenswerte müssen erst in Relation zum zugeordneten Zeitpunkt mittels der entsprechenden Inflationswerte vergleichbar gemacht werden. Der
- 5 Inflationsindex kann ebenfalls aus entsprechenden Datenbanken ausgelesen werden oder mittels Eingabeeinheiten dem System eingegeben werden. Der Inflationsindex für ein Land kann z.B. folgendermassen aussehen:

| Jahr | Inflationsindex (%) | Jährliche Inflationswert |
|------|---------------------|--------------------------|
| 1989 | 100                 | 1.000                    |
| 1990 | 105.042             | 1.050                    |
| 1991 | 112.920             | 1.075                    |
| 1992 | 121.429             | 1.075                    |
| 1993 | 128.676             | 1.060                    |
| 1994 | 135.496             | 1.053                    |
| 1995 | 142.678             | 1.053                    |
| 1996 | 148.813             | 1.043                    |
| 1997 | 153.277             | 1.030                    |
| 1998 | 157.109             | 1.025                    |
| 1999 | 163.236             | 1.039                    |
| 2000 | 171.398             | 1.050                    |
| 2001 | 177.740             | 1.037                    |
| 2002 | 185.738             | 1.045                    |

- Weitere Normierungsfaktoren sind ebenso vorstellbar, wie z.B.
- 10 regionale Abhängigkeiten etc. Werden Schadensfälle länderübergreifend verglichen bzw. extrapoliert, kommen entsprechende Länderabhängigkeiten dazu. Für den allgemeinen, nicht versicherungsspezifischen Fall kann sich die Normierung auch auf Abhängigkeiten, wie z.B. das mittlere Alter bei Populationen von Lebewesen, Natureinflüssen etc. etc. beziehen.

- 15 Zur automatisierten Bestimmung der Entwicklungswerte  $P_{i,K+2-i,f}, \dots, P_{i,K,f} = (Z_{i,K+2-i,f}, R_{i,K+2-i,f}), \dots, (Z_{i,K,f}, R_{i,K,f})$  umfasst das System und/oder Verfahren mindestens ein neuronales Netzwerk. Als neuronale Netzwerke

können z.B. konventionelle statische und/oder dynamische neuronale Netzwerke, wie beispielsweise feedforward (heteroassoziative) Netzwerke wie ein Perceptron oder ein Multi-Layer-Perceptron (MLP) gewählt werden, aber auch andere Netzwerkstrukturen, wie z.B. rekurrente Netzwerkstrukturen, sind vorstellbar. Die unterschiedliche Netzwerkstruktur der feedforward Netze im Gegensatz zu Netzwerke mit Rückkopplung (rekurrente Netzwerke) bestimmt, in welcher Art Informationen durch das Netzwerk verarbeitet werden. Im Falle eines statischen neuronalen Netzwerkes soll die Struktur die Nachbildung statischer Kennfelder mit ausreichender Approximationsgüte gewährleisten. Für dieses Ausführungsbeispiel seien als Beispiel Multi-Layer-Perceptrons gewählt. Ein MLP besteht aus mehreren Neuronenschichten mit mindestens einem Inputlayer und einem Outputlayer. Die Struktur ist strikt vorwärts gerichtet und gehört zur Gruppe der Feed-Forward-Netzen. Neuronale Netzwerke bilden ganz allgemein ein m-dimensionales Eingabesignal auf ein n-dimensionales Ausgabesignal ab. Die zu verarbeitende Information wird im hier betrachteten Feedforward-Netzwerk von einer Schicht mit Inputneuronen, dem Inputlayer, aufgenommen. Die Inputneuronen verarbeiten die Eingangssignale und geben sie über gewichtete Verbindungen, sog. Synapsen, an eine oder mehrere verdeckte Neuronenschichten, den Hiddenlayers, weiter. Von den Hiddenlayers wird das Signal ebenfalls mittels gewichteter Synapsen auf Neuronen eines Outputlayers übertragen, welcher ihrerseits das Ausgangssignal des neuronalen Netzwerkes generieren. In einem vorwärtsgerichteten, vollständig verbundenen MLP ist jedes Neuron eines bestimmten Layers mit allen Neuronen des nachfolgenden Layers verbunden. Die Wahl der Anzahl von Layers und Neuronen (Netzknoten) in einem bestimmten Layer ist wie üblich dem entsprechenden Problem anzupassen. Die einfachste Möglichkeit ist die ideale Netzstruktur empirisch zu ermitteln. Dabei ist zu beachten, dass bei einer zu gross gewählten Anzahl von Neuronen das Netzwerk anstatt zu lernen, rein abbildend wirkt, während es bei einer zu kleinen Anzahl von Neuronen zu Korrelationen der abgebildeten Parameter kommt. Anders ausgedrückt ist es so, dass wenn die Anzahl der Neuronen zu klein gewählt wird, dann kann die Funktion möglicherweise nicht dargestellt werden. Mit der Erhöhung der Anzahl der versteckten Neuronen steigt jedoch auch die Anzahl der unabhängigen Variablen in der Fehlerfunktion. Dies führt zu mehr lokalen Minima und der höheren Wahrscheinlichkeit in genau einer dieser Minima zu landen. Im

Spezialfall des Backpropagation kann dieses Problem z.B. mittels Simulated Annealing mindestens minimiert werden. Beim Simulated Annealing wird den Zuständen des Netzes eine Wahrscheinlichkeit zugeordnet. In der Analogie zum Kühlen von flüssiger Materie, aus denen Kristalle entstehen, wird eine  
 5 große Anfangstemperatur T gewählt. Diese wird nach und nach verkleinert, je kleiner umso langsamer. In der Analogie der Bildung von Kristallen aus Flüssigkeit geht man davon aus, dass falls man die Materie zu schnell abkühlen lässt, die Moleküle sich nicht gemäss der Gitterstruktur anordnen. Der Kristall wird unrein und an den betroffenen Stellen instabil. Um dies zu verhindern,  
 10 lässt man die Materie nun so langsam abkühlen, dass die Moleküle immer noch genügend Energie haben, um aus einen lokalen Minimum herauszuspringen. Bei den neuronalen Netzen wird nichts anderes gemacht: Es wird zusätzlich die Größe T in einer leicht veränderten Fehlerfunktion eingeführt. Diese konvergiert dann im Idealfall gegen ein globales Minimum.

15 Für die Anwendung zur Erfahrungstarifizierung haben sich bei MLP neuronale Netzwerke mit einer mindestens dreischichtigen Struktur als sinnvoll erwiesen. Das heisst, dass die Netzwerke mindestens einen Inputlayer, einen Hiddenlayer und einen Outputlayer umfassen. Innerhalb jedes Neurons finden die drei Verarbeitungsschritte Propagierung, Aktivierung und Ausgabe statt. Als  
 20 Ausgang des i-ten Neurons der k-ten Schicht ergibt sich

$$o_i^k = f_i^k \left( \sum_j w_{i,j}^k \cdot o_{i,j}^{k-1} + b_{i,j}^k \right)$$

wobei z.B. für k=2, als Bereich der Laufvariable j=1,2,...,N<sub>1</sub> gilt, mit N<sub>1</sub> wird die Anzahl der Neuronen des Layers k-1. w als Gewicht und b als Bias (Schwellwert) bezeichnet. Der Bias b kann je nach Anwendung für alle  
 25 Neuronen eines bestimmten Layers gleich oder unterschiedlich gewählt sein. Als Aktivierungsfunktion kann z.B. eine log-sigmoidale Funktion gewählt werden, wie

$$f_i^k(\xi) = \frac{1}{1 + e^{-\xi}}$$

Die Aktivierungsfunktion (oder Transferfunktion) wird in jedem  
 30 Neuron eingesetzt. Andere Aktivierungsfunktionen wie Tangentialfunktionen

etc. sind jedoch erfindungsgemäss ebenfalls möglich. Beim Backpropagation-Verfahren ist jedoch darauf zu achten, dass eine differenzierbare Aktivierungsfunktion, wie z.B. eine sigmoide Funktion, da dies Voraussetzung für das Verfahren ist. D.h. also z.B. binäre Aktivierungsfunktion wie z.B.

$$f(x) := \begin{cases} 1 & \text{falls } x > 0 \\ 0 & \text{falls } x \leq 0 \end{cases}$$

gehen für das Backpropagation Verfahren nicht. In den Neuronen der Ausgangsschicht werden die Ausgänge des letzten Hiddenlayers gewichtet aufsummiert. Die Aktivierungsfunktion des Outputlayers kann auch linear sein. Die Gesamtheit der Gewichtungen  $W_{i,j}^k$  und Bias  $B_{i,j}^k$  zusammengefasst in den Parameter- bzw. Wichtungsmatrizen bestimmen das Verhalten der neuronalen Netzstruktur

$$W^k = (w_{i,j}^k) \in \mathbb{R}^{N \cdot N_k}$$

Damit ergibt sich

$$o^k = B^k + W^k \cdot \left( 1 + e^{-(B^{k-1} + W^{k-1} \cdot u)} \right)^{-1}$$

Die Art und Weise, wie das Netzwerk ein Eingangssignal auf ein Ausgabesignal abbilden soll, d.h. die Bestimmung der gewünschten Gewichte und Bias des Netzwerkes, wird erreicht, indem das Netzwerk mittels Trainingsmuster trainiert wird. Der Satz der Trainingsmuster (Index  $\mu$ ) besteht aus dem Eingangssignal

$$Y^\mu = [y_1^\mu, y_2^\mu, \dots, y_{N_1}^\mu]$$

und einem Ausgangssignal

$$U^\mu = [u_1^\mu, u_2^\mu, \dots, u_{N_1}^\mu]$$

In diesem Ausführungsbeispiel mit der Erfahrungstarifizierung von Schadensfällen umfassen die Trainingsmuster die bekannten Ereignisse  $P_{i,f}$  mit

den bekannten Entwicklungswerten  $P_{ikf}$  für alle  $k$ ,  $f$  und  $i$ . Die Entwicklungswerte der zu extrapolierenden Ereignisse können dabei natürlich zum Training der neuronalen Netzwerke nicht verwendet werden, da zu ihnen der entsprechende Outputwert fehlt.

5 Zu Beginn des Lernvorgangs kann die Initialisierung der Gewichte der Hiddenlayers, in diesem Ausführungsbeispiel also der Neuronen, z.B. mit einer log-sigmoidale Aktivierungsfunktion, z.B. nach Nguyen-Widrow (D. Nguyen, B. Widrow, "Improving the Learning Speed of 2-Layer Neural Networks by Choosing Initial Values of Adaptive Weights", International Joint Conference  
10 of Neural Networks, vol 3, pp 21-26, July 1990) durchgeführt werden. Falls für die Neuronen des Outputlayers eine lineare Aktivierungsfunktion gewählt wurde, könne die Gewichte z.B. mittels eines symmetrischen Zufallsgenerators initialisiert werden. Zum Training des Netzwerkes können verschiedene Lernverfahren des Standes der Technik verwendet werden, wie z.B. das  
15 Backpropagation-Verfahren, Learning Vector Quantization, Radial Basis Funktion, Hopfield-Algorithmus oder Kohonen-Algorithmus etc. Die Aufgabe des Trainingsverfahrens besteht darin, die Synapsengewichte  $w_{ij}$  und Bias  $b_{ij}$  innerhalb der Wichtungsmatrix  $W$  bzw. der Biasmatrix  $B$  so zu bestimmen, dass die Eingabemuster  $Y^\mu$  auf die entsprechenden Ausgabemuster  $U^\mu$  abgebildet  
20 werden. Zur Beurteilung des Lernstadiums kann z.B. der absolute quadratische Fehler

$$Err = \frac{1}{2} \sum_{\mu=1}^P \sum_{\lambda=1}^m (u_{eff,\lambda}^\mu - u_{soll,\lambda}^\mu)^2 = \sum_{\mu=1}^P Err^\mu$$

verwendet werden. Der Fehler  $Err$  berücksichtigt dabei alle Muster  $P_{ikf}$  der Trainingsbasis, bei welchen die effektiven Ausgabesignale  $U_{eff}^\mu$  von den  
25 in der Trainingsbasis vorgegebenen Sollreaktionen  $U_{soll}^\mu$  zeigen. Für dieses Ausführungsbeispiel soll als Lernverfahren das Backpropagation-Verfahren gewählt werden. Das Backpropagation-Verfahren ist ein rekursives Verfahren zur Optimierung der Gewichtungsfaktoren  $w_{ij}$ . Bei jedem Lernschritt wird nach dem Zufallsprinzip ein Eingabemuster  $Y^\mu$  ausgewählt und durch das Netz propagiert  
30 (Forwardpropagation). Mittels der oben beschriebenen Fehlerfunktion  $Err$  wird aus dem vom Netzwerk generierten Ausgabesignal mit der in der Trainingsbasis vorgegebenen Sollreaktion  $U_{soll}^\mu$  der Fehler  $Err^\mu$  auf das



präsentierte Eingabemuster bestimmt. Die Änderungen der einzelnen Gewichte  $w_{ij}$  nach der Präsentation des  $\mu$ -ten Trainingsmusters sind dabei proportional zur negativen partiellen Ableitung des Fehlers  $Err^\mu$  nach dem Gewicht  $w_{ij}$  (sog. Gradientenabstiegsverfahren)

$$\Delta w_{i,j}^\mu \approx \frac{\partial E^\mu}{\partial w_{i,j}}$$

Mit Hilfe der Kettenregel können aus der partiellen Ableitung die als Backpropagation-Regel bekannten Adaptionvorschriften für die Elemente der Wichtungsmatrix bei der Präsentation des  $\mu$ -ten Trainingsmusters hergeleitet werden.

$$\Delta w_{i,j}^\mu \equiv s \cdot \delta_i^\mu \cdot u_{eff,j}^\mu$$

mit

$$\delta_i^\mu = f^1(\xi_i^\mu) \cdot (u_{soll,i}^\mu - u_{eff,i}^\mu)$$

für den Outputlayer bzw.

$$\delta_i^\mu = f^1(\xi_i^\mu) \cdot \sum_k^K \delta_k^\mu w_{k,i}$$

für die Hiddenlayers. Der Fehler wird dabei beginnend mit dem Outputlayer in umgekehrter Richtung durch das Netzwerk propagiert (Backpropagation) und gewissermassen nach dem Verursacherprinzip auf die einzelnen Neuronen aufgeteilt. Der Proportionalitätsfaktor  $s$  wird als Lernfaktor bezeichnet. Während der Trainingsphase wird einem neuronalen Netzwerk eine begrenzte Anzahl an Trainingsmustern präsentiert, welche die zu erlernende Abbildung ausreichend genau charakterisieren. In diesem Ausführungsbeispiel mit der Erfahrungstarifizierung von Schadensfällen können die Trainingsmuster alle bekannten Ereignisse  $P_{i,f}$  mit den bekannten Entwicklungswerten  $P_{i,kf}$  für alle  $k, f$  und  $i$  umfassen. Aber auch eine Auswahl aus den bekannten Ereignissen  $P_{i,f}$  ist vorstellbar. Wird dem Netzwerk anschliessend ein Eingabesignal präsentiert, welches nicht exakt mit den

Mustern der Trainingsbasis übereinstimmt, so inter- bzw. extrapoliert das Netzwerk im Rahmen der erlernten Abbildungsfunktion zwischen den Trainingsmustern. Diese Eigenschaft wird als Generalisierungsfähigkeit der Netzwerke bezeichnet. Es ist charakteristisch für neuronale Netzwerke, dass

5 neuronale Netzwerke eine gute Fehlertoleranz besitzen. Dies ist ein weiterer Vorteil gegenüber den Systemen des Standes der Technik. Da neuronale Netzwerke eine Vielzahl von (teilweise redundanten) Eingangssignalen auf das/die gewünschten Ausgabesignal/e abbilden, erweisen sich die Netzwerke als robust gegenüber Ausfall einzelner Eingangssignale bzw. gegenüber

10 Signalrauschen. Eine weitere interessante Eigenschaft neuronaler Netzwerke ist deren Lernfähigkeit. Prinzipiell ist es daher möglich, ein einmal trainiertes System während des Betriebs permanent/periodisch nachlernen oder anpassen zu lassen, was ebenfalls ein Vorteil gegenüber den Systemen des Standes der Technik ist. Für das Lernverfahren können natürlich auch andere

15 Verfahren verwendet werden, wie z.B. ein Verfahren nach Levenberg-Marquardt (D. Marquardt, "An Algorithm for least square estimation of non-linear Parameters", J.Soc.Ind.Appl.Math, pp 431-441, 1963 sowie M.T. Hagan, M.B.Menjaj, "Training Feedforward Networks with the Marquardt Algorithm", IEEE-Transactions on Neural Networks, Vol 5, Nr 6, pp 989-993, November

20 1994). Das Levenberg-Marquardt-Verfahren ist eine Kombination der Gradient-Methode und des Newton-Verfahrens und hat den Vorteil, dass es schneller konvergiert als das oben erwähnte Backpropagation Verfahren, jedoch eine höhere Speicherkapazität während der Trainingsphase benötigt.

In dem Ausführungsbeispiel werden zur Bestimmung der

25 Entwicklungswerte  $P_{i,K-(i-j)+1,f}$  für jedes Anfangsjahr  $i$  iterativ  $(i-1)$  neuronale Netzwerke  $N_{i,j}$  erzeugt.  $j$  gibt für ein bestimmtes Anfangsjahr  $i$  die Anzahl Iterationen an mit  $j=1, \dots, (i-1)$ . Damit werden für das  $i$ -te Anfangsjahr  $i-1$  neuronale Netzwerke  $N_{i,j}$  erzeugt. Das neuronale Netzwerk  $N_{i,j+1}$  hängt dabei rekursiv vom neuronalen Netzwerk  $N_{i,j}$  ab. Zum Gewichten, d.h. zum Trainieren,

30 eines bestimmten neuronalen Netzwerkes  $N_{i,j}$  können z.B. alle Entwicklungswerte  $P_{p,q,f}$  mit  $p=1, \dots, (i-1)$  und  $q=1, \dots, K-(i-j)$  der Ereignisse oder Schadensfälle  $P_{pq}$  verwendet werden. Eine eingeschränkte Auswahl kann je nach Anwendung jedoch ebenfalls sinnvoll sein. Die Daten der Ereignisse  $P_{pq}$  können z.B. wie erwähnt aus einer Datenbank ausgelesen werden und über

eine Recheneinheit dem System präsentiert werden. Ein errechneter Entwicklungswert  $P_{i,k,f}$  kann z.B. dem entsprechenden Ereignis  $P_{i,f}$  eines Anfangsjahres  $i$  zugeordnet werden und selbst zur Bestimmung des nächsten Entwicklungswertes (z.B.  $P_{i,k+1,f}$ ) dem System präsentiert werden (Fig. 1 bis 6)  
 5 oder die Zuordnung findet erst nach Ende der Bestimmung aller gesuchten Entwicklungswerte  $P$  statt (Fig. 7).

Im ersten Fall (Fig. 1 bis 6) werden, wie beschrieben, einem bestimmten Ereignis  $P_{i,f}$  eines Anfangsjahres  $i$  Entwicklungswerte  $P_{i,k,f}$  mit Entwicklungsjahr  $k=1,\dots,K$  zugeordnet, wobei für die Anfangsjahre  $i = 1,\dots,K$  und  
 10  $K$  das letzt bekannte Entwicklungsjahr ist. Für das erste Anfangsjahr  $i=1$  sind alle Entwicklungswerte  $P_{1,k,f}$  bekannt. Für jedes Anfangsjahr  $i=2,\dots,K$  mittels Iterationen  $j=1,\dots,(i-1)$  wird bei jeder Iteration  $j$  in einem ersten Schritt ein neuronales Netzwerk  $N_{i,j}$  mit einem Inputlayer mit  $K-(i-j)$  Inputsegmenten und einem Outputlayer erzeugt. Jedes Inputsegment umfasst mindestens ein  
 15 Inputneuron bzw. mindestens so viele Inputneurone, um das Inputsignal für einen Entwicklungswert  $P_{i,k,f}$  zu erhalten. Die neuronalen Netzwerke werden automatisch durch das System erzeugt und können hardwaremässig oder softwaremässig realisiert sein. In einem zweiten Schritt wird das neuronale Netzwerk  $N_{i,j}$  mit den verfügbaren Ereignissen  $E_{i,f}$  aller Anfangsjahre  $m=1,\dots,(i-1)$   
 20 mittels der Entwicklungswerte  $P_{m,1..K-(i-j),f}$  als Input und  $P_{m,1..K-(i-j)+1,f}$  als Output gewichtet. In einem dritten Schritt werden mittels des neuronalen Netzwerkes  $N_{i,j}$  die Outputwerte  $O_{i,f}$  für alle Ereignisse  $P_{i,f}$  des Anfangsjahres  $i$  bestimmt, wobei der Outputwert  $O_{i,f}$  dem Entwicklungswert  $P_{i,K-(i-j)+1,f}$  des Ereignisses  $P_{i,f}$  zugeordnet wird und wobei das neuronale Netzwerk  $N_{i,j}$  rekursiv vom  
 25 neuronalen Netzwerk  $N_{i,j+1}$  abhängt. Fig. 1 zeigt die Trainings- bzw. Präsentationsphase eines neuronalen Netzwerkes zur Bestimmung des Ereigniswertes  $P_{2,5,f}$  eines Ereignisses  $P_f$  in einer oberen  $5 \times 5$  Matrix, d.h. bei  $K=5$ . Die gestrichelte Linie T gibt die Trainingsphase an und die ausgezogene Linie R die Bestimmungsphase nach dem Lernen. Figur 2 zeigt das Gleiche für  
 30 das dritte Anfangsjahr zur Bestimmung von  $P_{3,4,f}$  ( $B_{34}$ ) und Figur 3 zur Bestimmung von  $P_{3,5,f}$ . Figur 4 zeigt nur die Trainingsphase zur Bestimmung von  $P_{3,4,f}$  und  $P_{3,5,f}$ , wobei die erzeugten Werte  $P_{3,4,f}$  ( $B_{34}$ ) zum Trainieren des Netzwerkes zur Bestimmung von  $P_{3,5,f}$  verwendet werden.  $A_{ij}$  gibt in den Figuren die bekannten Werte an, während  $B_{ij}$  mittels der Netzwerke bestimmte

Werte anzeigt. Figur 5 zeigt die rekursive Erzeugung der neuronalen Netzwerke zur Bestimmung der Werte in Zeile 3 einer 5x5 Matrix, wobei i-1 Netzwerke erzeugt werden, also 2. Figur 6 zeigt dagegen die rekursive Erzeugung der neuronalen Netzwerke zur Bestimmung der Werte in Zeile 3 einer 5x5 Matrix, wobei wiederum i-1 Netzwerke erzeugt werden, also 4.

Es ist wichtig darauf hinzuweisen, dass als Ausführungsbeispiel die Zuordnung der mittels des Systems erzeugten Ereigniswerte  $B_{ij}$  auch erst nach der Bestimmung aller gesuchten Entwicklungswerte  $P$  stattfinden kann. Die neu bestimmten Werte stehen dann nicht als Inputwerte zur Bestimmung weiterer Ereigniswerte zur Verfügung. Figur 7 zeigt ein solches Verfahren, wobei die Trainingsbasis auf die bekannten Ereigniswerte  $A_{ij}$  eingeschränkt ist. Anders ausgedrückt können die neuronalen Netzwerk  $N_{i,j}$  für gleiche  $j$  identisch sein, wobei für ein Anfangszeitintervall  $i+1$  das neuronale Netzwerk  $N_{i+1,j=i}$  erzeugt wird und alle anderen neuronalen Netzwerke  $N_{i+1,j<i}$  Netzwerken früheren Anfangszeitintervallen entsprechen. D.h. ein Netzwerk, das einmal zur Berechnung eines bestimmten Ereigniswertes  $P_{ij}$  erzeugt wurde, wird für alle Ereigniswerte mit einem Anfangsjahr  $a>i$  für die Werte  $P_{ij}$  mit gleichem  $j$  weiterverwendet.

Unterschiedliche neuronalen Netzwerke können im Falle der hier diskutierten Versicherungsfälle z.B. basierend auf unterschiedlichen Daten trainiert werden. Z.B. können die Netzwerke basierend auf den bezahlten Forderungen, basierend auf den angefallenen Forderungen, basierend auf den bezahlten und noch ausstehenden Forderungen (Reserven) und/oder basierend auf den bezahlten und angefallenen Forderungen trainiert werden. Das beste neuronale Netzwerk für jeden Fall kann z.B. mittels Minimierung des absoluten mittleren Fehlers der vorhergesagten Werte und den wirklichen Werte bestimmt werden. Z.B. lässt sich das Verhältnis des mittleren Fehlers zum mittleren vorhergesagten Wert (der bekannten Forderungen) auf die vorhergesagten Werte der modellierten Werte anwenden, um den Fehler zu erhalten. Für den Fall, dass die vorhergesagten Werte der vorgängigen Anfangsjahre zur Berechnung der folgenden Anfangsjahre mitverwendet wird, muss der Fehler natürlich entsprechend kumuliert werden. Dies kann z.B.

erreicht werden, indem die Quadratwurzel der Summe der Quadrate der individuellen Fehler jedes Modells verwendet wird.

Um eine weitere Abschätzung der Qualität bzw. des Trainingstandes der neuronalen Netzwerke zu erhalten, können z.B. die vorhergesagten Werte auch mittels der genannten Pareto-Verteilung gefittet werden. Diese Abschätzung kann ebenfalls verwendet werden, um z.B. das beste neuronale Netzwerk von mit (wie im letzten Abschnitt beschrieben) unterschiedlichen Datensets trainierten neuronalen Netzwerken (z.B. bezahlten Forderungen, ausstehenden Forderungen etc.) zu bestimmen. Damit folgt mit der Pareto-Verteilung

$$\chi^2 = \sum \left( \frac{O(i) - T(i)}{E(i)} \right)^2$$

mit

$$T(i) = Th(1 - P(i))^{(-1/\alpha)}$$

wobei  $\alpha$  der Fit-Parameter, Th der Schwellparameter (Threshold Value),  $T(i)$  der theoretische Wert der i-ten Zahlungsforderung,  $O(i)$  der beobachtete Wert der i-ten Zahlungsforderung,  $E(i)$  ist der Fehler der i-ten Zahlungsforderung und  $P(i)$  ist die kummulierte Wahrscheinlichkeit der i-ten Zahlungsforderung mit

$$P(1) = \left( \frac{1}{2n} \right)$$

und

$$P(i+1) = P(i) + \frac{1}{n}$$

und n die Anzahl Zahlungsforderungen. Für das hier beschriebene Ausführungsbeispiel wurde der Fehler der Systeme basierend auf den vorgeschlagenen neuronalen Netzwerken mit dem Chain-Ladder-Verfahren anhand von Kraftfahrzeugversicherungsdaten verglichen. Die Netzwerke

wurden einmal mit den bezahlten Zahlungsforderungen und einmal mit den angefallenen Zahlungsforderungen verglichen. Um die Daten zu Vergleichen, wurden die einzelnen Werte in den Entwicklungsjahren kumuliert. Der direkte Vergleich zeigte für die gewählten Beispielesdaten pro 1000 folgende Resultate

| Anfangs<br>-jahr | System basierend auf neuronalen<br>Netzwerken |   | Chain-Ladder-Verfahren                        |   |
|------------------|---|---|---|---|
|                  | Bezahlte<br>Forderungen<br>(kumulierte Werte) | Angefallenen<br>Forderungen<br>(kumulierte Werte) | Bezahlte<br>Forderungen<br>(kumulierte Werte) | Angefallenen<br>Forderungen<br>(kumulierte Werte) |
| 1996             | 369.795 ± 5.333                               | 371.551 ± 6.929                                   | 387.796 ± n/a                                 | 389.512 ± n/a                                     |
| 1997             | 769.711 ± 6.562                               | 789.997 ± 8.430                                   | 812.304 ± 0.313                               | 853.017 ± 15.704                                  |
| 1998             | 953.353 ± 40.505                              | 953.353 ± 30.977                                  | 1099.710 ± 6.522                              | 1042.908 ± 32.551                                 |
| 1999             | 1142.874 ± 84.947                             | 1440.038 ± 47.390                                 | 1052.683 ± 138.221                            | 1385.249 ± 74.813                                 |
| 2000             | 864.628 ± 99.970                              | 1390.540 ± 73.507                                 | 1129.850 ± 261.254                            | 1285.956 ± 112.668                                |
| 2001             | 213.330 ± 72.382                              | 288.890 ± 80.617                                  | 600.419 ± 407.718                             | 1148.555 ± 439.112                                |

- 5 Der hier gezeigte Fehler entspricht der Standardabweichung, d.h. dem  $\sigma_1$ -Fehler, der angegebenen Werte. Besonders für spätere Anfangsjahre, d.h. Anfangsjahre mit grösserem  $i$  zeigt das System basierend auf neuronalen Netzwerken bei der Bestimmung der Werte einen klaren Vorteil gegenüber den Verfahren des Standes der Technik, indem die Fehler im wesentlichen stabil
- 10 bleiben. Dies ist beim Stand der Technik nicht der Fall, da der Fehler dort für zunehmende  $i$  nicht proportional zunimmt. Für grössere Anfangsjahre  $i$  zeigt sich eine deutliche Abweichung in der Höhe der kumulierten Werte zwischen den Chain Ladder Werten und denen, welche mit dem erfindungsgemässen Verfahren erhalten wurden. Diese Abweichung beruht auf der Tatsache, dass
- 15 beim Chain Ladder Verfahren zusätzlich die IBNYR (Incurred But Not Yet Reported) Schäden berücksichtigt werden. Die IBNYR Schäden müssten zu den obengezeigten Werten des erfindungsgemässen Verfahrens addiert werden. Z.B. kann zur Berechnung von Portefeuillereserven die IBNYR Schäden mittels einer separaten Entwicklung (z.B. Chain Ladder) berücksichtigt
- 20 werden. Bei der Reservierung von Einzelschäden oder bei der Ermittlung von Schadenshöhenverteilungen spielen die IBNYR Schäden jedoch keine Rolle.

## Ansprüche

1. Computergestütztes System zur automatisierten Erfahrungstarifizierung und/oder Schadensreservierung, wobei ein bestimmtes Ereignis  $P_{i,f}$  eines Anfangszeitintervalles  $i$  Entwicklungswerte  $P_{ikf}$  der

5    Entwicklungsintervalle  $k=1,...,K$  umfasst, wobei  $K$  das letzte bekannte Entwicklungsintervall ist mit  $i=1,...,K$  und wobei für die Ereignisse  $P_{1,f}$  alle Entwicklungswerte  $P_{1kf}$  bekannt sind, dadurch gekennzeichnet,

        dass das System zur automatisierten Bestimmung der Entwicklungswerte  $P_{i,K+2-i,f}, \dots, P_{i,K,f}$  mindestens ein neuronales Netzwerk

10    umfasst.
2. Computergestütztes System nach Anspruch 1, dadurch gekennzeichnet, dass für die Ereignisse das Anfangszeitintervall einem Anfangsjahr entspricht und die Entwicklungsintervalle Entwicklungsjahren entsprechen.
- 15        3. Computergestütztes System nach einem der Ansprüche 1 oder 2, dadurch gekennzeichnet, dass das System zur Bestimmung der Entwicklungswerte  $P_{i,K+2-i,f}, \dots, P_{i,K,f}$  eines Ereignis  $P_{i,f}$  ( $i-1$ ) iterativ erzeugte neuronale Netzwerke  $N_{i,j}$  für jedes Anfangszeitintervall  $i$  umfasst mit  $j=1,...,(i-1)$ , wobei das neuronale Netzwerk  $N_{i,j+1}$  rekursiv vom neuronalen Netzwerk  $N_{i,j}$

20    abhängt.
4. Computergestütztes System nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, dass Trainingswerte zum Gewichten eines bestimmten neuronalen Netzwerkes  $N_{i,j}$  die Entwicklungswerte  $P_{p,q,f}$  mit  $p=1,...,(i-1)$  und  $q=1,...,K-(i-j)$  umfassen.
- 25        5. Computergestütztes System nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, dass die neuronalen Netzwerk  $N_{i,j}$  für gleiche  $j$  identisch sind, wobei für ein Anfangszeitintervall  $i+1$  das neuronale Netzwerk  $N_{i+1,j=i}$  erzeugt wird und alle anderen neuronalen Netzwerke  $N_{i+1,j<i}$  Netzwerken früheren Anfangszeitintervallen entsprechen.

6. Computergestütztes System nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, dass das System zusätzlich Ereignisse  $P_{i,f}$  mit Anfangszeitintervall  $i < 1$  umfasst, wobei für die Ereignisse  $P_{i < 1, f}$  alle Entwicklungswerte  $P_{i < 1, k, f}$  bekannt sind.

5 7. Computergestütztes System nach einem der Ansprüche 1 bis 6, dadurch gekennzeichnet, dass das System mindestens einen Normierungsfaktor umfasst, mittels welchem die Entwicklungswerte  $P_{i, k, f}$  der verschiedenen Ereignis  $P_{i, f}$  entsprechend ihrem Anfangszeitintervall normierbar sind.

10 8. Computergestütztes Verfahren zur automatisierten Erfahrungstarifizierung und/oder Schadensreservierung, wobei einem bestimmten Ereignis  $P_{i, f}$  eines Anfangszeitintervalles  $i$  Entwicklungswerte  $P_{i, k, f}$  mit Entwicklungsintervallen  $k = 1, \dots, K$  zugeordnet werden, wobei  $K$  das letzte bekannte Entwicklungsintervall ist mit  $i = 1, \dots, K$  und wobei für die Ereignisse  $P_{1, f}$   
15 alle Entwicklungswerte  $P_{1, k, f}$  bekannt sind, dadurch gekennzeichnet,

dass zur Bestimmung der Entwicklungswerte  $P_{i, K+2-i, f}, \dots, P_{i, K, f}$  mindestens ein neuronales Netzwerk verwendet wird.

9. Computergestütztes Verfahren nach Anspruch 8, dadurch gekennzeichnet, dass für die Ereignisse das Anfangszeitintervall einem  
20 Anfangsjahr zugeordnet wird und die Entwicklungsintervalle Entwicklungsjahren zugeordnet werden.

10. Computergestütztes Verfahren nach einem der Ansprüche 8 oder 9, dadurch gekennzeichnet, dass zur Bestimmung der Entwicklungswerte  $P_{i, K-(i-j)+1, f}$  iterativ  $(i-1)$  neuronale Netzwerke  $N_{i, j}$  für jeden Anfangszeitintervall  $i$   
25 erzeugt werden mit  $j = 1, \dots, (i-1)$ , wobei das neuronale Netzwerk  $N_{i, j+1}$  rekursiv vom neuronalen Netzwerk  $N_{i, j}$  abhängt.

11. Computergestütztes Verfahren nach einem der Ansprüche 8 bis 10, dadurch gekennzeichnet, dass zum Gewichten eines bestimmten



neuronalen Netzwerkes  $N_{i,j}$  die Entwicklungswerte  $P_{p,q,f}$  mit  $p=1, \dots, (i-1)$  und  $q=1, \dots, K-(i-j)$  verwendet werden.

12. Computergestütztes Verfahren nach einem der Ansprüche 8 bis 10, dadurch gekennzeichnet, dass die neuronalen Netzwerk  $N_{i,j}$  für gleiche  $j$  identisch trainiert werden, wobei für ein Anfangszeitintervall  $i+1$  das neuronale Netzwerk  $N_{i+1,j=i}$  erzeugt wird und alle anderen neuronalen Netzwerke  $N_{i+1,j<i}$  von früheren Anfangszeitintervall übernommen werden.

13. Computergestütztes Verfahren nach einem der Ansprüche 8 bis 12, dadurch gekennzeichnet, dass zur Bestimmung zusätzlich Ereignisse  $P_{i,f}$  mit Anfangszeitintervall  $i < 1$  verwendet werden, wobei für die Ereignisse  $P_{i < 1, f}$  alle Entwicklungswerte  $P_{i < 1, k, f}$  bekannt sind.

14. Computergestütztes Verfahren nach einem der Ansprüche 8 bis 13, dadurch gekennzeichnet, dass mittels mindestens einen Normierungsfaktor die Entwicklungswerte  $P_{i,k,f}$  der verschiedenen Ereignis  $P_{i,f}$  entsprechend ihrem Anfangszeitintervall normiert werden.

15. Computergestütztes Verfahren zur automatisierten Erfahrungstarifizierung und/oder Schadensreservierung, wobei einem bestimmten Ereignis  $P_{i,f}$  eines Anfangszeitintervall  $i$  Entwicklungswerte  $P_{i,k,f}$  mit Entwicklungsintervallen  $k=1, \dots, K$  zugeordnet abgespeichert werden, wobei  $i = 1, \dots, K$  und  $K$  das letzte bekannte Entwicklungsintervallen ist und wobei für das erste Anfangszeitintervall alle Entwicklungswerte  $P_{1,k,f}$  bekannt sind, dadurch gekennzeichnet,

25 dass für jedes Anfangszeitintervall  $i=2, \dots, K$  mittels Iterationen  $j=1, \dots, (i-1)$  bei jeder Iteration  $j$  in einem ersten Schritt ein neuronales Netzwerk  $N_{i,j}$  mit einem Inputlayer mit  $K-(i-j)$  Inputsegmenten und einem Outputlayer erzeugt wird, wobei jedes Inputsegment mindestens ein Inputneuron umfasst und einem Entwicklungswert  $P_{i,k,f}$  zugeordnet wird,

dass in einem zweiten Schritt das neuronale Netzwerk  $N_{i,j}$  mit den verfügbaren Ereignissen  $P_{i,f}$  aller Anfangszeitintervalle  $m=1, \dots, (i-1)$  mittels der

Entwicklungswerte  $P_{m,1..K-(l-j),f}$  als Input und  $P_{m,1..K-(l-j)+1,f}$  als Output gewichtet wird, und

dass in einem dritten Schritt mittels des neuronalen Netzwerkes  $N_{i,j}$  die Outputwerte  $O_{i,f}$  für alle Ereignisse  $P_{i,f}$  des Anfangsjahres  $i$  bestimmt werden, wobei der Outputwert  $O_{i,f}$  dem Entwicklungswert  $P_{i,K-(l-j)+1,f}$  des Ereignisses  $P_{i,f}$  zugeordnet wird und wobei das neuronale Netzwerk  $N_{i,j}$  rekursiv vom neuronalen Netzwerk  $N_{i,j+1}$  abhängt.

16. Computergestütztes Verfahren nach Anspruch 15, dadurch gekennzeichnet, dass für die Ereignisse das Anfangszeitintervall einem Anfangsjahr zugeordnet wird und die Entwicklungsintervalle Entwicklungsjahren zugeordnet werden.

17. System von neuronalen Netzwerken, welche neuronalen Netzwerke  $N_i$  jeweils einen Inputlayer mit mindestens einem Inputsegment und einen Outputlayer umfassen, wobei der Input- und Outputlayer eine Vielzahl von Neuronen umfasst, die gewichtet miteinander verbunden sind, dadurch gekennzeichnet,

dass die neuronalen Netzwerke  $N_i$  mittels einer Recheneinheit software- und/oder hardwaremässig iterativ erzeugbar sind, wobei ein neuronales Netzwerk  $N_{i+1}$  rekursiv vom neuronalen Netzwerk  $N_i$  abhängt und jedes Netzwerk  $N_{i+1}$  jeweils ein Inputsegment mehr als das Netzwerk  $N_i$  umfasst,

dass jedes neuronale Netzwerk  $N_i$  beginnend beim neuronalen Netzwerk  $N_1$  mittels eines Minimierungsmoduls durch Minimierung eines lokal propagierten Fehlers trainierbar ist, und

dass das rekursive System von neuronalen Netzwerken mittels einem Minimierungsmodul durch Minimierung eines global propagierten Fehlers basierend auf den lokalen Fehlern der neuronalen Netzwerke  $N_i$  trainierbar ist.

18. System von neuronalen Netzwerken nach Anspruch 17, dadurch gekennzeichnet, dass der Outputlayer des neuronalen Netzwerkes  $N_i$  mit mindestens einem Inputsegment des Inputlayers des neuronalen Netzwerkes  $N_{i+1}$  zugeordnet verbunden ist.

5 19. Computerprogrammprodukt, welches ein computerlesbares Medium mit darin enthaltenen Computerprogrammcodemitteln zur Steuerung eines oder mehrerer Prozessoren eines computer-basierten Systems zur automatisierten Erfahrungstarifizierung und/oder Schadensreservierung umfasst, wobei einem bestimmten Ereignis  $P_{i,f}$  eines Anfangszeitintervalles  $i$   
 10 Entwicklungswerte  $P_{i,k,f}$  mit Entwicklungsintervall  $k=1, \dots, K$  zugeordnet abgespeichert umfasst, wobei  $i = 1, \dots, K$  und  $K$  das letzte bekannte Entwicklungsintervall ist und wobei für das erste Anfangszeitintervall  $i=1$  alle Entwicklungswerte  $P_{1,k,f}$  bekannt sind, dadurch gekennzeichnet,

15 dass mittels des Computerprogrammproduktes mindestens ein neuronales Netzwerk softwaremässig generierbar ist und zur Bestimmung der Entwicklungswerte  $P_{i,K+2-f}, \dots, P_{i,K,f}$  verwendbar ist.

20 20. Computerprogrammprodukt nach Anspruch 19, dadurch gekennzeichnet, dass für die Ereignisse des Anfangszeitintervall einem Anfangsjahr zugeordnet ist und die Entwicklungsintervalle Entwicklungsjahren zugeordnet sind.

21. Computerprogrammprodukt nach einem der Ansprüche 19 oder 20, dadurch gekennzeichnet, dass zur Bestimmung der Entwicklungswerte  $P_{i,K-(i-j)+1,f}$  für jedes Anfangszeitintervall  $i$  mittels des Computerprogrammprodukt iterativ  $(i-1)$  neuronale Netzwerke  $N_{i,j}$  erzeugbar sind mit  $j=1, \dots, (i-1)$ , wobei das  
 25 neuronale Netzwerk  $N_{i,j+1}$  rekursiv vom neuronalen Netzwerk  $N_{i,j}$  abhängt.

22. Computerprogrammprodukt nach einem der Ansprüche 19 bis 21, dadurch gekennzeichnet, dass zum Gewichten eines bestimmten neuronalen Netzwerkes  $N_{i,j}$  mittels des Computerprogrammprodukt die Entwicklungswerte  $P_{p,q,f}$  mit  $p=1, \dots, (i-1)$  und  $q=1, \dots, K-(i-j)$  aus einer Datenbank  
 30 auslesbar sind.

23. Computerprogrammprodukt nach einem der Ansprüche 19 bis 21, dadurch gekennzeichnet, dass beim Computerprogrammprodukt die neuronalen Netzwerk  $N_{i,j}$  für gleiche  $j$  identisch trainiert werden, wobei mittels des Computerprogrammprodukt für ein Anfangszeitintervall  $i+1$  das neuronale  
5 Netzwerk  $N_{i+1,j=i}$  erzeugt wird und alle anderen neuronalen Netzwerke  $N_{i+1,j<i}$  von früheren Anfangszeitintervall übernommen werden.

24. Computerprogrammprodukt nach einem der Ansprüche 19 bis 23, dadurch gekennzeichnet, dass die Datenbank zusätzlich Ereignisse  $P_{i,f}$  mit Anfangszeitintervall  $i<1$  abgespeichert umfasst, wobei für die Ereignisse  $P_{i<1,f}$   
10 alle Entwicklungswerte  $P_{i<1,k,f}$  bekannt sind.

25. Computerprogrammprodukt nach einem der Ansprüche 19 bis 24, dadurch gekennzeichnet, dass das Computerprogrammprodukt mindestens einen Normierungsfaktor umfasst, mittels welchem die Entwicklungswerte  $P_{i,k,f}$  der verschiedenen Ereignis  $P_{i,f}$  entsprechend ihrem Anfangszeitintervall  
15 normierbar sind.

26. Computerprogrammprodukt, welches in den internen Speicher eines digitalen Computers ladbar ist und Softwarecodeabschnitte umfasst, mit denen die Schritte gemäss einem der Ansprüche 8 bis 16 durchführbar sind, wenn das Produkt auf einem Computer läuft, wobei die neuronalen Netzwerke  
20 softwaremässig und/oder hardwaremässig generierbar sind.

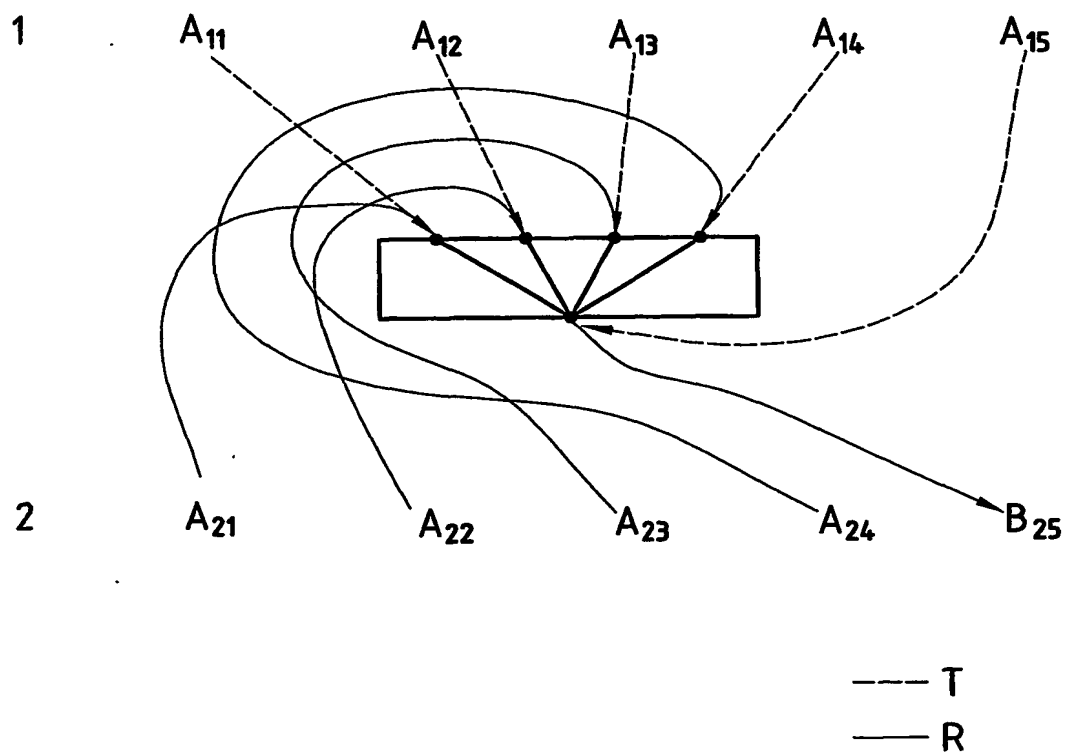


FIG. 1

2/7

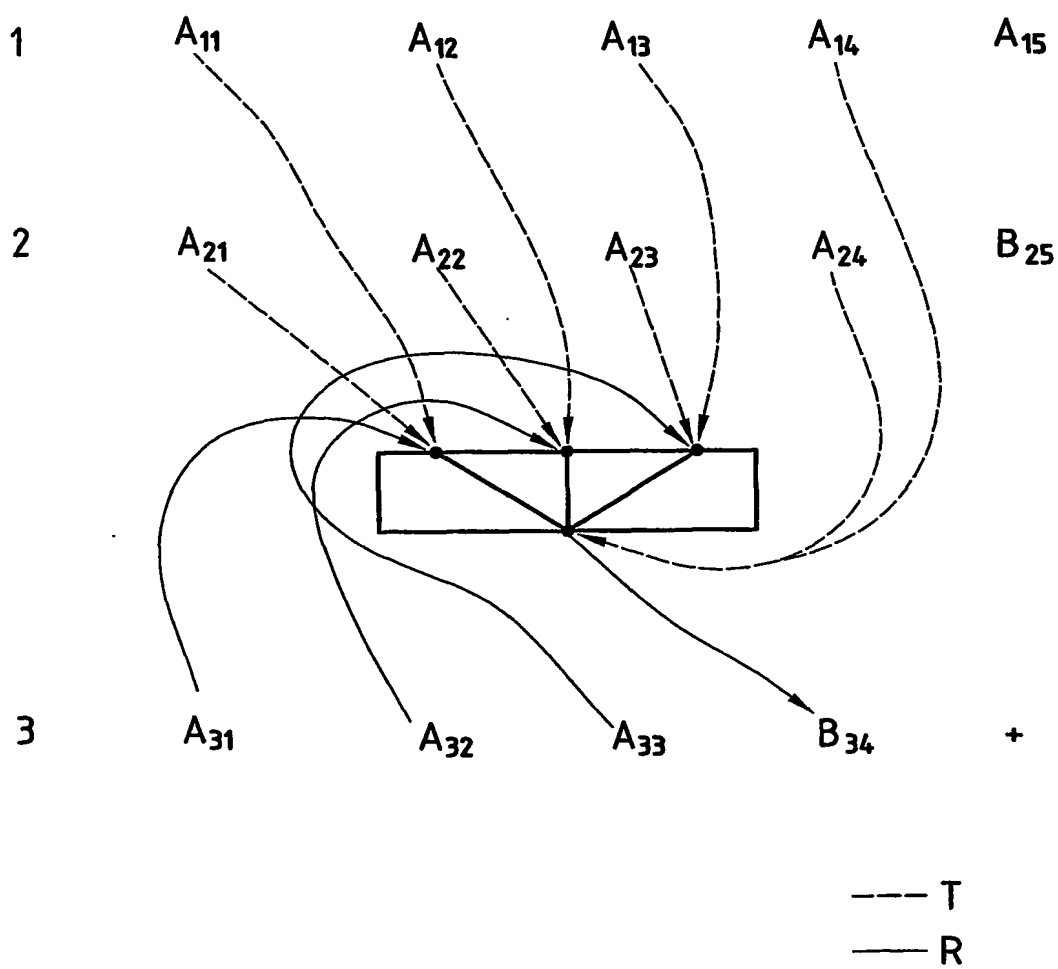


FIG. 2

3/7

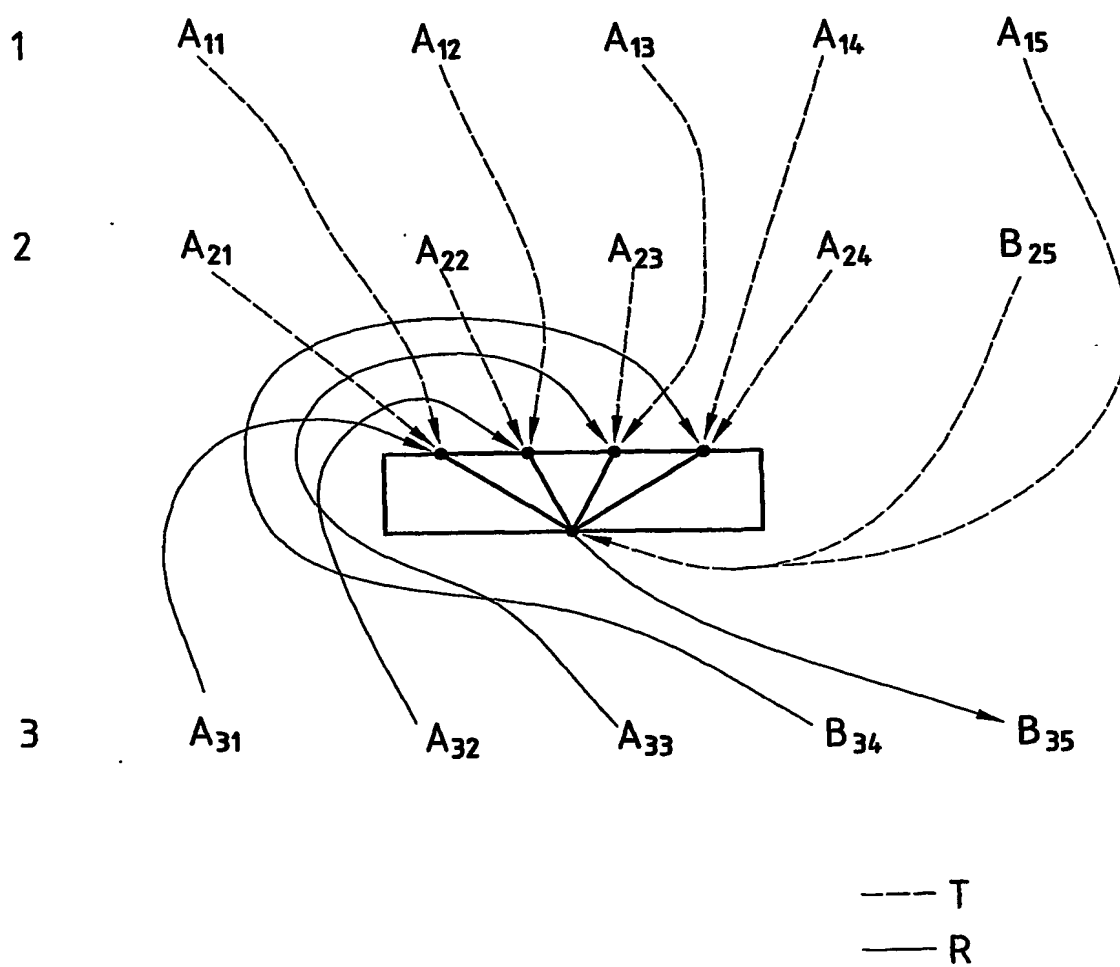


FIG. 3

4/7

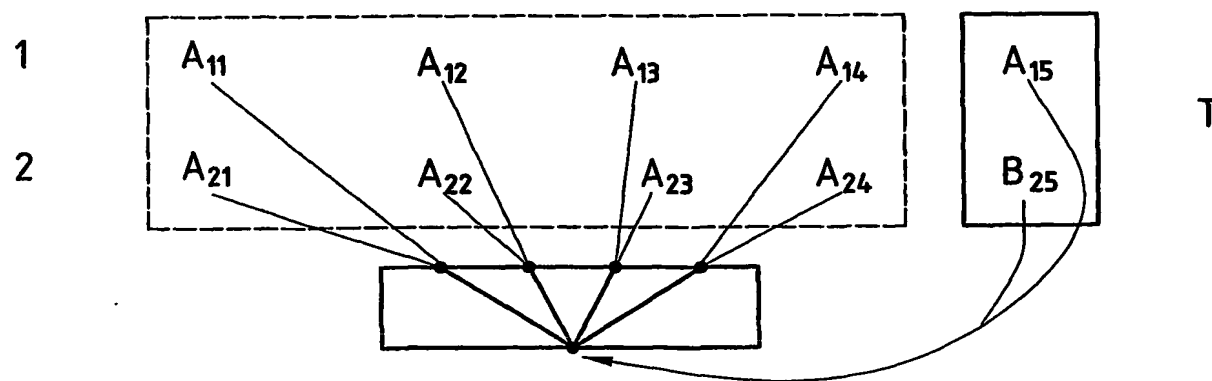
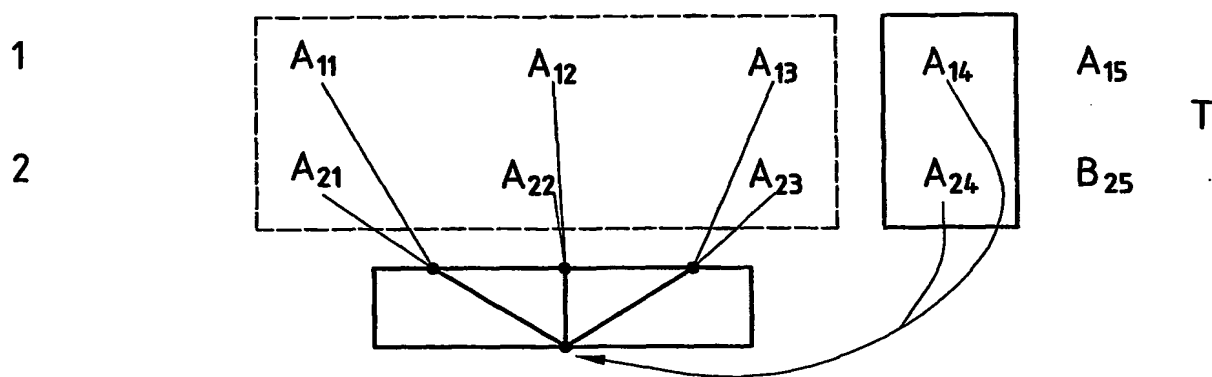


FIG. 4



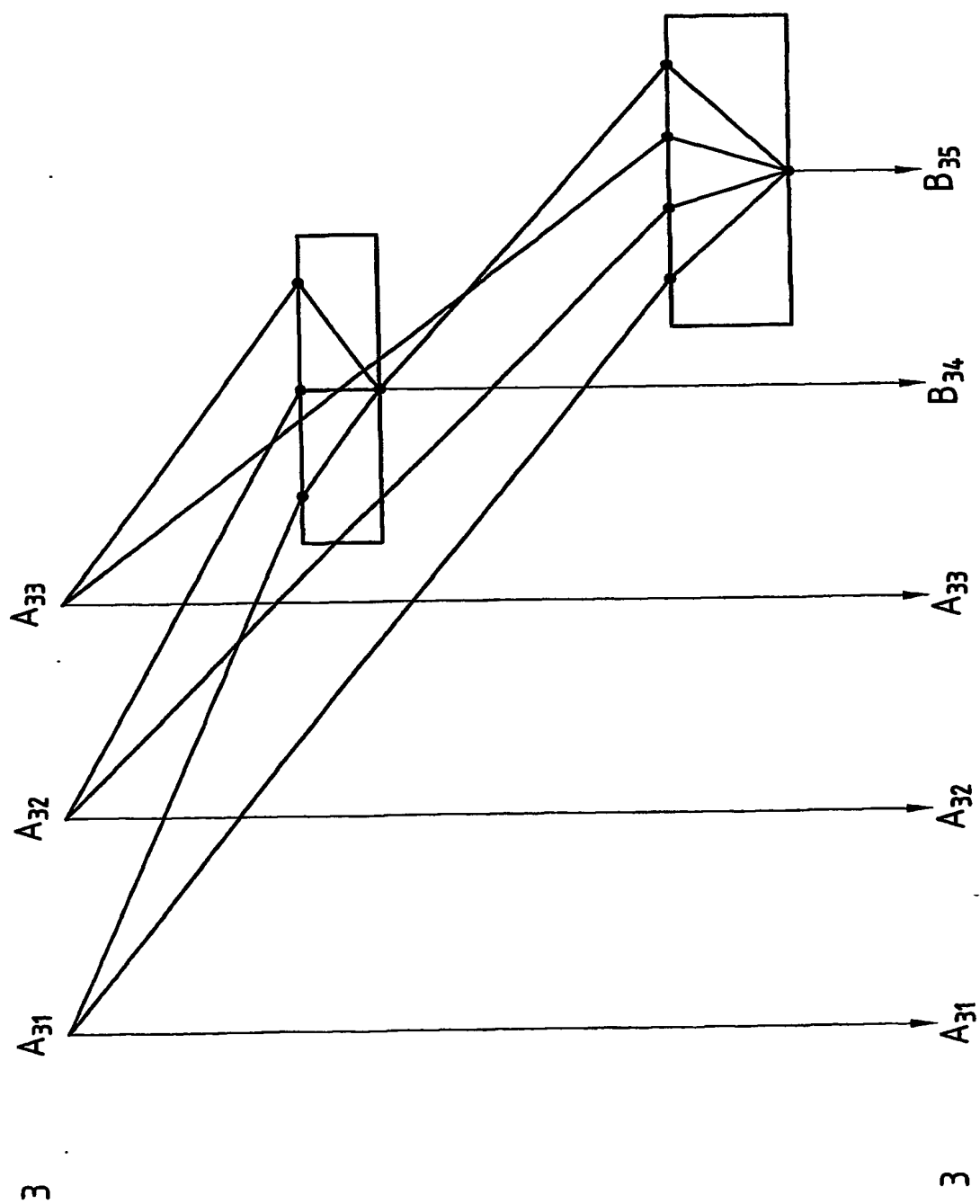
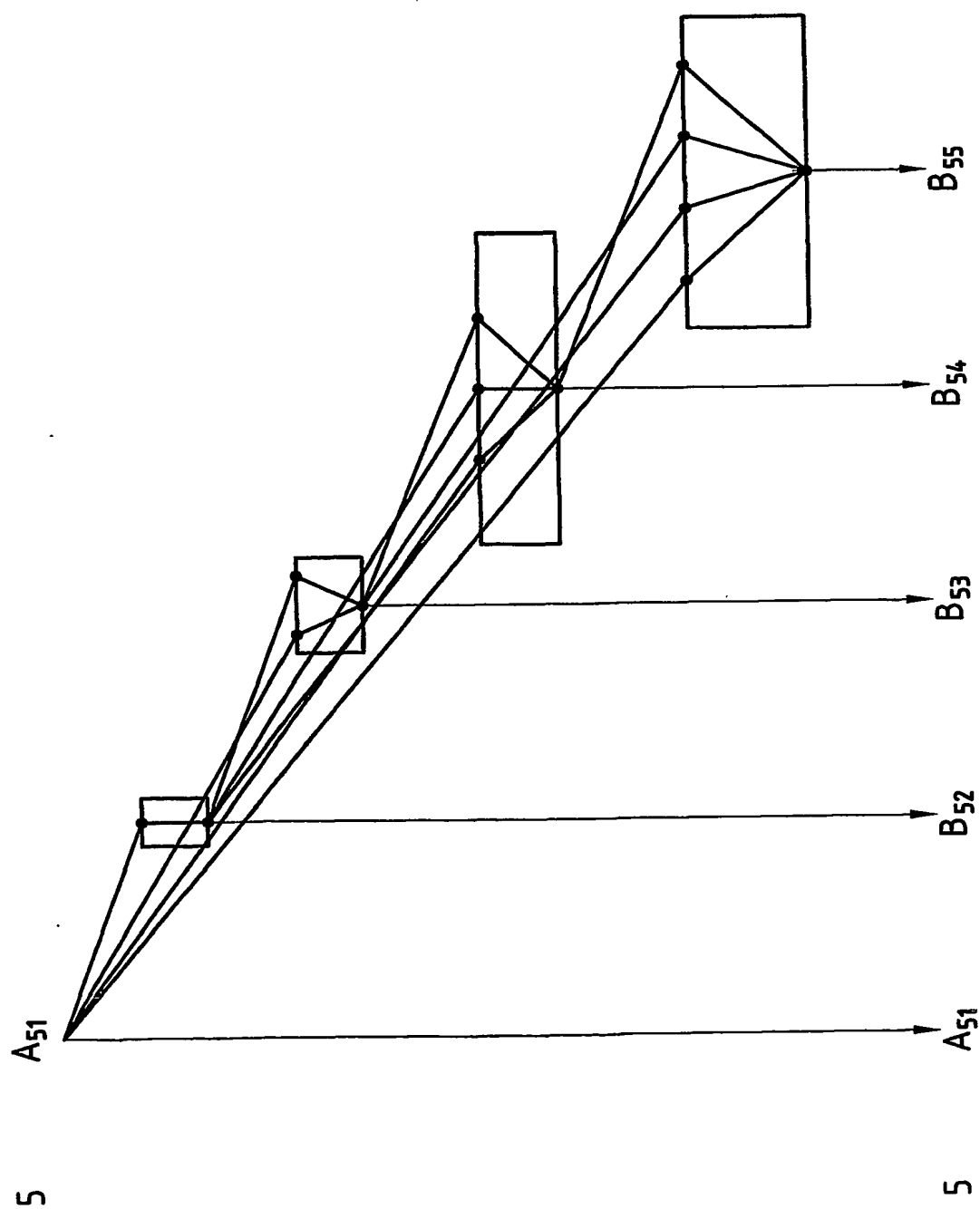


FIG. 6



7/7

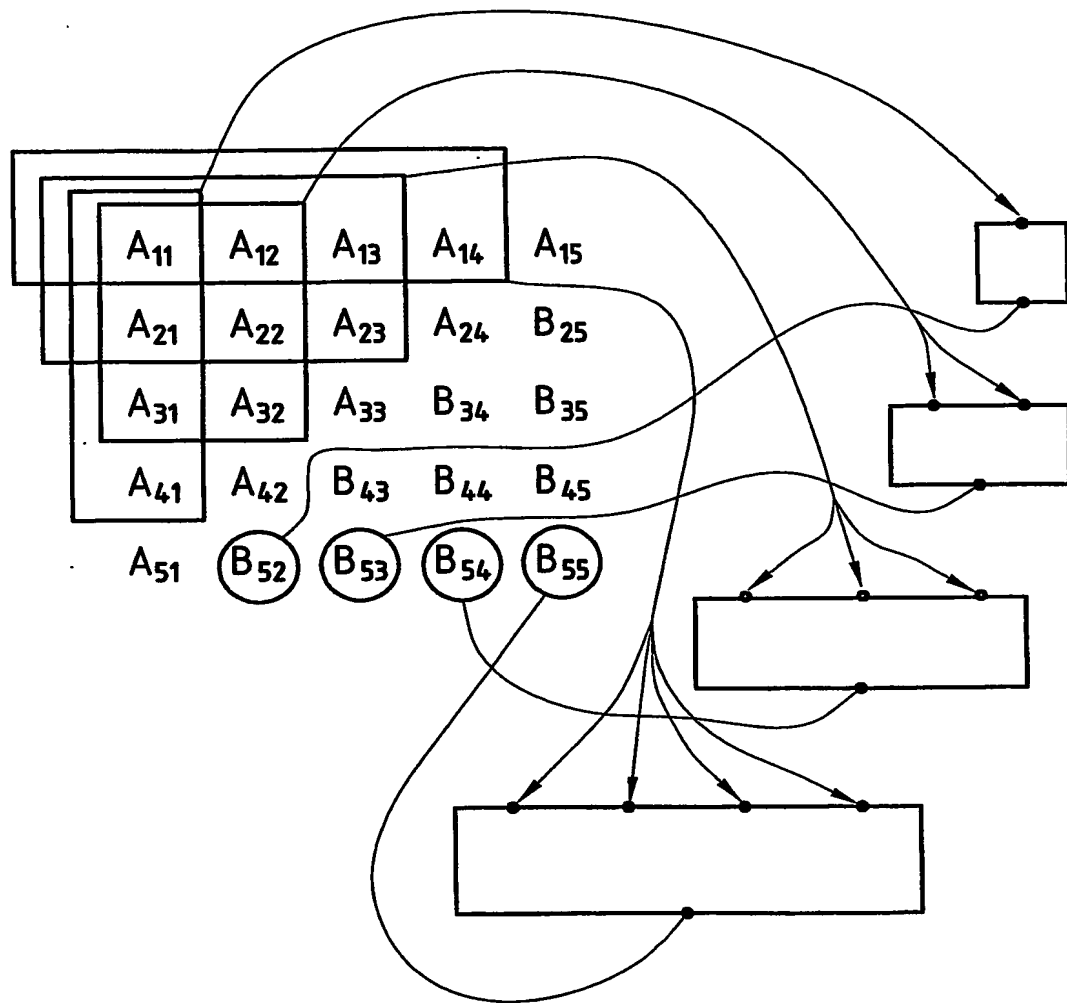


FIG. 7

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/CH 03/00612

**A. CLASSIFICATION OF SUBJECT MATTER**  
IPC 7 G06N3/04 G06N3/08 G06F17/60

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 G06F G06N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ, INSPEC

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages  | Relevant to claim No. |
|------------|---|-----------------------|
| A          | US 5 987 444 A (LO JAMES TING-HO)<br>16 November 1999 (1999-11-16)<br>abstract<br>column 23, line 60 - column 34, line 22 | 1-26                  |
| A          | US 5 761 442 A (BARR DEAN S ET AL)<br>2 June 1998 (1998-06-02)<br>abstract<br>column 5, line 66 - column 14, line 11      | 1-26                  |
| A          | WO 02/47026 A (HARBECK NADIA ; KATES<br>RONALD E (DE)) 13 June 2002 (2002-06-13)  |                       |

☐ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

### \* Special categories of cited documents:

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

- \*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- \*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- \*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- \*Z\* document member of the same patent family

Date of the actual completion of the international search

7 June 2004

Date of mailing of the international search report

16/06/2004

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Jacobs, P

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/CH 03/00612

| Patent document<br>cited in search report |   | Publication<br>date | Patent family<br>member(s) | Publication<br>date |
|---|---|---------------------|----------------------------|---------------------|
| US 5987444                                | A | 16-11-1999          | NONE                       |                     |
| US 5761442                                | A | 02-06-1998          | NONE                       |                     |
| WO 0247026                                | A | 13-06-2002          | AU 1608002 A               | 18-06-2002          |
|   |   |                     | WO 0247026 A2              | 13-06-2002          |
|   |   |                     | EP 1384199 A2              | 28-01-2004          |
|   |   |                     | US 2004073096 A1           | 15-04-2004          |

## INTERNATIONAL RESEARCH REPORT

Internationales Aktenzeichen  
PCT/CH 03/00612A. KLASSIFIZIERUNG DES ANMELDUNGSGEGENSTANDES  
IPK 7 G06N3/04 G06N3/08 G06F17/60

Nach der Internationalen Patentklassifikation (IPK) oder nach der nationalen Klassifikation und der IPK

## B. RESEARCHIERTE GEBIETE

Recherchierter Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole)  
IPK 7 G06F G06N

Recherchierte aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen

Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe)

EPO-Internal, WPI Data, PAJ, INSPEC

## C. ALS WESENTLICH ANGESEHENE UNTERLAGEN

| Kategorie* | Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile                                  | Betr. Anspruch Nr. |
|------------|---|--------------------|
| A          | US 5 987 444 A (LO JAMES TING-HO)<br>16. November 1999 (1999-11-16)<br>Zusammenfassung<br>Spalte 23, Zeile 60 - Spalte 34, Zeile 22 | 1-26               |
| A          | US 5 761 442 A (BARR DEAN S ET AL)<br>2. Juni 1998 (1998-06-02)<br>Zusammenfassung<br>Spalte 5, Zeile 66 - Spalte 14, Zeile 11      | 1-26               |
| A          | WO 02/47026 A (HARBECK NADIA ; KATES<br>RONALD E (DE)) 13. Juni 2002 (2002-06-13)   |                    |

☐ Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen☒ Siehe Anhang Patentfamilie

\* Besondere Kategorien von angegebenen Veröffentlichungen :

\*A\* Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist

\*E\* älteres Dokument, das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist

\*L\* Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt)

\*O\* Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht

\*P\* Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist

\*T\* Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist

\*X\* Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erfinderischer Tätigkeit beruhend betrachtet werden

\*Y\* Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erfinderischer Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist

\*Z\* Veröffentlichung, die Mitglied derselben Patentfamilie ist

Datum des Abschlusses der internationalen Recherche

7. Juni 2004

Absenddatum des internationalen Recherchenberichts

16/06/2004

Name und Postanschrift der Internationalen Recherchenbehörde  
Europäisches Patentamt, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax (+31-70) 340-3016

Bevollmächtigter Bediensteter

Jacobs, P

# INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen  
PCT/CH 03/00612

| Im Recherchenbericht<br>angeführtes Patentdokument |   | Datum der<br>Veröffentlichung | Mitglied(er) der<br>Patentfamilie |               | Datum der<br>Veröffentlichung |
|--|---|-------------------------------|-----------------------------------|---------------|-------------------------------|
| US 5987444   | A | 16-11-1999                    | KEINE                             |               |                               |
| US 5761442   | A | 02-06-1998                    | KEINE                             |               |                               |
| WO 0247026   | A | 13-06-2002                    | AU                                | 1608002 A     | 18-06-2002                    |
|  |   |                               | WO                                | 0247026 A2    | 13-06-2002                    |
|  |   |                               | EP                                | 1384199 A2    | 28-01-2004                    |
|  |   |                               | US                                | 2004073096 A1 | 15-04-2004                    |